DOCUMENT RESUME

ED 231 339 IR 010 736

AUTHOR	Mills, Michael I.
TITLE	A Study of the Human Response to Pictorial Representations on Telidon. Telidon Behavioural Research 3.
INSTITUTION	Montreal Univ. (Quebec).
SPONS AGENCY	Department of Communications, Ottawa (Ontario).
PUB DATE	[80]
CONTRACT	09SU.36001-0-3165
NOTE	160p.
PUB TYPE	Information Analyses (070) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC07 Plus Postage.
DESCRIPTORS	*Computer Graphics; Design Requirements; *Information Networks; Intermode Differences; Media Research; *Perception; *Pictorial Stimuli; Realism; *Videotex
IDENTIFIERS	*Telidon System

ABSTRACT

The role pictures and graphics can play in making Telidon, a Canadian home information retrieval videotex service, a viable communications medium is discussed. This report concentrates on contributions of theories of perception and cognition to knowledge of how different kinds of visual images can most effectively be used on Telidon. The issues of whether some ideas are best translated into images while others are more suited to words and how pictures and words might complement each other in communication information on videotex are first addressed. The need for pictorial realism in portraying objects in videotex displays is then examined. Because studies on how people identify objects in pictures suggest that, in some cases, line-drawings, cartoons, and caricatures can more effectively convey information than realistic photographs, the implications of these findings for how pictures of objects are stored in the data base are explained. Finally, the report examines how abstract images and schemas can facilitate learning and problem-solving on videotex, especially how dynamic images can be used on Telidon. Also included are a short introduction to picture perception psychology by Adam Gopnik, and a four-page supplementary bibliography. (LMM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED231339

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

● Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

A STUDY OF THE HUMAN RESPONSE TO PICTORIAL REPRESENTATIONS ON TELIDON

Telidon Behavioural Research 3

BY

Michael I. Mills

UNIVERSITE DE MONTREAL

PREPARED FOR:

The Behavioural Research and Evaluation Group
Information Technology Branch
Department of Communications
Ottawa, Canada

Contract: 09SU.36001-0-3165

IR010736

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
P. Racine

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EXECUTIVE SUMMARY

The use of pictures and graphics on Telidon, while potentially very rewarding, can entail certain costs to information-providers, system operators and users which, ultimately, can affect the overall acceptability of the service. In view of these costs, which can vary with the type and complexity of image, it is necessary to ask what benefits derive from the use of visual images on Telidon. The goals of this report are (1) to review current research in the areas of human cognition and perception on how people understand and use visual images; and (2) to examine the implications of this work for the role of pictures and graphics on Telidon. The report is divided into three main sections, each containing a number of themes which, while not exhaustive, are seen as relevant to the problem of visual imagery on Telidon.

Chapter I examines the relative merits of images and words in the communication process by considering the question: Can a picture ever substitute for a verbal statement? Answering this question reveals some fundamental differences between pictures and words and shows that a picture cannot easily match the descriptive functions of language. In addition, some theoretical notions are introduced which help one understand why pictures pose difficult problems in applied areas relevant to Telidon: e.g., the automatic classification and retrieval of images by computer and the development of "iconic indexes." Finally, it is proposed that the real challenge to information-providers is not to find ways of replacing words with pictures on Telidon, but to understand how a picture can serve as a "conceptual base" facilitating the comprehension of text and how text can guide the processing of a picture.

Chapter II asks if there will be a problem of "pictorial literacy" on Telidon, or can anyone who can see the visual world, be able to see the world through a picture? In addition, research is reviewed which suggests that pictures departing from photographic realism -- outline-drawings, cartoons and caricatures -- may be more easily and quickly identified than photographs of the real-thing. This research on the perception of line-drawings and caricatures suggests a number of tradeoffs (e.g., transmission-time versus pictorial realism) which can exploit Telidon's strength as a pictorial medium: namely, its ability to generate high-quality line-graphics.

Chapter III examines the use of simple graphic imagery to communicate complex ideas. It looks at what is involved in learning from and using maps and raises the problem of individual differences in people's ability to understand certain kinds of graphic imagery. In addition, this chapter explores why simple graphic schemas (e.g., Venn diagrams) can aid thinking and learning. It is proposed that the usefulness of abstract graphic imagery for learning is related to its ability to give visible shape to abstract properties and relations not directly observable in the physical appearances of objects and events. Finally, it is suggested that dynamic graphic schemas showing "transformations" will be more effective aids to productive thinking and problem-solving than static graphics and that it may be possible to specify dynamic transformations on Telidon, economically, by using successive still-samples without providing the illusion of continuous motion.

Finally, Appendix A provides the interested reader with an "Introduction to the Psychology of Picture Perception" prepared especially for this report by Adam Gopnik.

The views and interpretations advanced in this document are my own and not necessarily those of the Department of Communications.

ACKNOWLEDGEMENTS

I would like to thank Dr. Dorothy Phillips of the Department of Communications, the scientific authority for this project, for her guidance in helping to clarify the ideas contained in this report. Thanks are also due to other members of the Behavioural Research Group, Eric Lee, Bill Treurniet, Paul Muter and Tom Whelan, with whom I've had useful discussions. I would especially like to express my appreciation to Adam Gopnik, my research assistant on this project, who not only conducted the bulk of the literature search, but who participated integrally in the shaping of the ideas. Adam is the author of the "Introduction to Picture Perception" found in Appendix A.

TABLE OF CONTENTS

EXECUTIVE SUMMARY.. ii

ACKNOWLEDGEMENTS... v

INTRODUCTION... 1

- 0.1. Why Study the Role of Visual Images on Telidon?.. 1
- 0.1.1. Building Pictures through Drawing "Primitives"... 2
- 0.2. Three Themes...................................... 7

CHAPTER I - WILL A PICTURE BE WORTH 10^3 WORDS ON TELIDON?... 12

- 1.0. A Popular Argument about Pictures and Words...... 13
- 1.1. Can a Picture Ever Substitute for a Verbal Statement?.. 15
- 1.1.1. Pictures Are Worth 10^3 Words, But Which Ones?.... 18
- 1.2. Multiple Meanings in Pictures: A "Schema" Approach to Perception.......................... 19
- 1.2.1. Schemas and Transformations...................... 21
- 1.3. Picture Descriptions for Machines................ 27
- 1.4. Can Pictures Serve As Indexes?................... 30
- 1.5. The Importance of Context........................ 40
- 1.5.1. What Words Do For Pictures....................... 40
- 1.5.2. What Pictures Can Do For Words................... 44
- 1.5.3. Words and Images Interact........................ 47
- 1.6. Summary and Conclusion........................... 52

CHAPTER II - READING PICTURES: IS REALISM ALWAYS BEST?..... 56

- 2.0. Do We Learn to Read Pictures?.................... 56
- 2.1. The "Language" of Lines.......................... 63
- 2.2. The Time/Realism Tradeoff: Implications of Using Outline-Pictures on Telidon................ 67
- 2.3. Photos, Line-Drawings, Cartoons, and Caricatures. 75
- 2.3.1. Cartoons Give External Form to Internal Schemes.. 76
- 2.4. Summary and Conclusion........................... 88

CHAPTER III - IMAGES FOR THINKING: GRAPHIC SCHEMES FOR COMPLEX IDEAS.................................. 92

- 3.0. Learning from Maps............................... 93
- 3.0.1. Individual Differences in Learning Maps.......... 97
- 3.0.2. Movie Maps....................................... 99
- 3.1. Graphic Schemes for Thinking..................... 103
- 3.1.2. Dynamic Graphics................................. 107
- 3.1.3. Motion versus Sequence........................... 114
- 3.2. Summary and Conclusion........................... 115

CONCLUSION.. 119

REFERENCES.. 124

APPENDIX A: An Introduction to the Psychology of Picture Perception by Adam Gopnik...................... 130

APPENDIX B: Supplementary Bibliography..................... 146

LIST OF FIGURES

Figure:

0.1 - The numbers indicate the order in which the parts of the image appear on the user's display.. 4

1.0 - "The cat is on the mat"........................ 16

1.1 - A cartoon-strip analogy......................... 24

1.2 - A Steinberg cartoon............................ 28

1.3 - Iconic Indexes Based on Those Used During the Montreal Olympics............................... 38

1.4 - Droodles: a) A midget playing a trombone in a telephone booth; b) An early bird who caught a very strong worm............................. 42

1.5 - Picture from the Bransford and Johnson (1972) experiment..................................... 45

1.6 - Examples of the four kinds of drawings that accompanied the word pairs presented in the Wollen et al. experiment....................... 48

1.7 - A single cartoon representing two different events: (a) A foot stepping on a bug (side-view); and (b) A car colliding with a pole (top-view)................................ 50

2.0 - Drawing (a) can be a picture of a cube since the converging lines obey the law of linear perspective; Drawing (b) cannot, strictly speaking, be a picture of a cube since its receding edges are parallel.................... 59

2.1 - The "meaning" of a Line......................... 64

2.2 - An outline drawing of a house traced from a color photograph............................... 70

2.3 - A picture matrix for "breadth-first" search.... 74

2.4 - Three of the four types of pictures used in the Ryan and Schwartz experiment. The fourth picture was an actual photograph.................. 77

2.5 - Two possible projections of a cube............... 80

2.6 - Tracings of photographs (a) and caricatures (b) of Richard Nixon................................. 83

2.7 - Sample photographs and caricatures used in the Perkins and Hagen (1980) experiment.......... 87

3.0 - A subway map.................................... 94

3.1 - The actual layout of an apartment (a) and a 12 year-old girl's reproduction from memory (b).. 96

3.2 - A visual metaphor for "good marriage/bad marriage" 105

3.3 - Geometric problem-solving........................ 109

3.4 - Still-frames from the "dynamic symbols" film..... 112

INTRODUCTION

0.1. Why Study the Role of Visual Images on Telidon?

Telidon is the name used to refer to a type of information-retrieval service developed by the Canadian Department of Communications, by which a person, using their home television set and a key pad, can access banks of information stored in a central computer. The generic term for such an interactive information-retrieval service is "Videotex."[1] The kinds of information a person might retrieve using a Videotex system are many and range from such standards as news, weather and stocks, to potentially more specialized kinds of interactions such as: searching for a home, games, how do to do one's taxes, how to prepare a gourmet meal or even learning mathematics.

Telidon not only provides the user with information in the form of the printed word, but also enables the user to retrieve information in the form of pictures and graphics. Now while visual images of all sorts abound on television, movies, magazines, billboard, text-books, and so on, and have many uses, from education to advertising, the role of the visual image on Telidon is somewhat problematical. But to understand why this is so, we will first have to take a brief look into how pictures and text are created on Telidon -- at least the current version.

---

[1] A detailed description of the nature of Videotex, which compares Telidon's technical characteristics with other Videox systems in Europe and the U.S., is given in Woolfe (1980). For a definition of Telidon's terminal and transmission characteristics, see Bown, O'Brien, Sawchuck and Storey (1978, 1979).

(My description of the page-creation process which follows will be very general since there exist full descriptions elsewhere of how pictures, text and graphics are created, stored and retrieved on Telidon. C.f. Bown et al., 1979)

0.1.1. Building Pictures through Drawing "Primitives."

Let us imagine that you are an information-provider, not just a consumer, and that you want to create some pages of information for Telidon on some topic, say, how to sail a boat. You are immediately faced with a number of sticky questions concerning how to design pages of information which will allow the user to quickly grasp the important ideas you want to get across. Some of these questions concern devising an appropriate data-structure which captures the important conceptual relations among the ideas. Other questions concern the choice of a communications medium in which you will present your ideas to the user. Given that you want to provide information about sailboats, for example, what aspects of this information should be presented to the user in the form of text, graphics or some combination of text and graphics? Creating a page of text on Telidon which describes aspects of how to sail a boat poses no special problems. Of course, you must decide what to say and the best way to say it, and this may demand some creativity, but once done, you merely use the special page creation terminal to type in the text -- making decisions about page formating, colour, and so on. Generating alphanumeric characters then is a relatively straightforward process on Telidon.

But let's say you want to include a simple line-drawing of a sailboat as one of the pages of information. (See Figure 0.1). To make such a drawing on the existing Telidon system you would use a special graphics language called Picture Description Instructions (PDI's). The PDI's are a set of drawing commands which allow you, by using a light-pen or key presses, to specify the sizes and shapes of various parts of an image and where they will appear on the screen. The PDI's include such drawing "primitives" as POINT, LINE, ARC POLYGON, etc., and include commands for choosing from a range of colours. So, for example, to draw the simple sailboat scene in Figure 0.1, you might begin by using the POLYGON command to define the vertices of the hull-shape on the graphics screen; the computer then connects up the points of the vertices and the hull-shape appears; next, to draw the sun, you would use the ARC command to define the starting point, a center point and an end point which, once interpreted by the computer, draws a circular "sun" on the screen. The other parts of the picture -- the mast, the horizon-line, etc., would be created in the same way by selecting and executing commands through the PDI's. It is important to note, for the purpose of this report, that the order in which the parts of the picture were described, is the order which the user will see when the drawing is retrieved. So, for instance, if you had drawn the parts of the image in the order indicated by the numbers in Figure 0.1 (first the hull, the sun, then the horizon-line, etc.), this is the sequence in which these segments would be stored in the image description and is the order in which the parts of the sailboat would appear on the user's screen -- not all at once -- but part-by-part at the time of retrieval. For a simple drawing

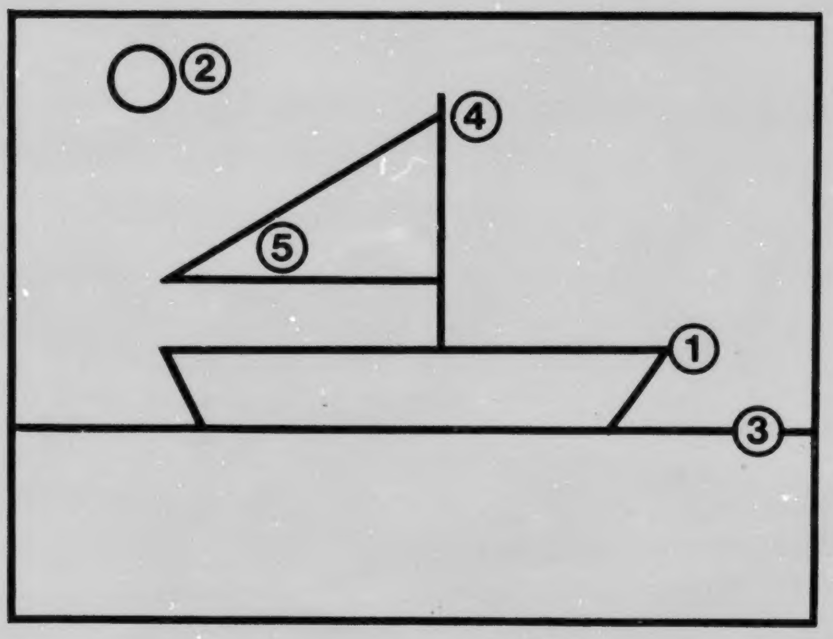


Figure 0.1. The numbers indicate the order in which the parts of the image appear on the user's display.

such as Figure 0.1, the segments would appear in rapid sequence, say within several seconds. For more complex images, more PDI commands must be used to describe the image and the longer it would take for the picture to build up on the user's terminal.

It should be noted that to store and display photographic images, it would be possible to by-pass the drawing primitives and input a picture using an "alphaphotographic" or "bit map" display mode in which -- instead of using the geometric primitives to describe a graphic image -- a photograph would be encoded as a large number of individual dots or picture elements. During the retrieval of the picture these dots would be mapped directly onto the user's television screen in a "slow-scan" or facsimile mode whereby the image "fills up" the screen gradually, line-by-line from top-to-bottom or from bottom-to-top. This "bit-map" approach allows the transmission of photographic-like images and by-passes the PDI's but, as we shall see in later chapters, this alternative has penalties of transmission delays (it would take on the order of several minutes for a photographic image to fill up an entire screen using the bit-map mode) and storage requirements.[1]

The use of PDI's to describe images has important advantages over alternative approaches to graphics such as those adopted by certain European Videotex systems: for example, PDI's provide for higher quality graphics and allow the description of images in a format which can adapt to changes in display hardware -- e.g., terminals with different resolution capacities, once armed with the appropriate decoder, could interpret PDI's (c.f., Bown et al., 1978). But the important point, for present purposes, is to realize that providing information in the form of a visual image for Telidon -- either through PDI's or the "bit-map" mode -- is not exactly like the normal way of drawing pictures on paper or photographing images for television, but can involve certain costs and tradeoffs some of which we can summarize as follows:

---

[1] Time of display depends on the speed of transmission line. Lines with higher speed than those commonly used at present would remove the problem.

(1) For an information provider, the decision to embody an idea as a visual image, as opposed to a textual description, may entail costs in terms of the time and effort required to conceive of, and construct an image using the PDI's. In general, it appears to be more complicated to make a picture than to input a textual description for Telidon, the time and effort rising as a function of image complexity.

(2) While a visual image may provide valuable information for a a user, there may also be certain costs associated with receiving pictures. One type of cost is the amount of time a user spends waiting for the PDI's to build up, part-by-part, on the viewer's screen. The more parts the image contains, the longer the wait. In general, though, simple graphic images, defined by PDI's, will take much less time to appear (seconds) than transmitting photographic images through the bit-map mode (minutes). Another kind of cost may be the cognitive effort spent trying to understand a poorly conceived picture or graphic or, conversely, the time wasted on a flashy, complex image which contributes nothing to one's comprehension, but only serves to distract. There are many potential costs

of this kind associated with the ad-hoc use of images on Telidon.

(3) Finally, using images can cost the system operator or information-provider in terms of memory space required to store an image description in the central computer. These memory costs rise with the complexity of the graphic, but less memory is consumed by storing graphic images described as PDI's than by storing graphics or photographs as a large number of individual picture elements.

0.2. Three Themes

These generalizations make it easier to understand why one might want to study the role of pictorial imagery on Telidon. The use of pictures and graphics on Telidon, while potentially very exciting, brings with it certain costs to system operators, information-providers and, ultimately, to users, which can affect the overall acceptance of the service. In light of these costs, which can vary with the type and complexity of image, it becomes necessary to ask what benefits accrue from the use of various kinds of images on Telidon? And this means getting a clearer idea of how visual images function; how pictures contribute to the communication process in general. Toward this end, the present report can be viewed as an attempt to answer to two basic questions: (1) What can current research in the areas of perception and cognition tell us about how people understand and use visual images? and (2) What implications do theories of how people process and use pictorial information have for the use of pictorial and graphic imagery on Telidon? In other words, what can theories of pictorial comprehension tell us about how to exploit Telidon's strong points as a graphics medium?

Now getting answers to such broad questions is obviously a huge undertaking. At some point, in order for the answers to be useful, one must narrow down the many possible questions raised by the subject of pictorial communication to a more limited, but still relevant, set of themes. In reviewing the literature, three such themes have emerged. And, although still quite broad in scope, these themes at least provide an organizing framework for the report. We can summarize these three themes, which provide our chapter headings, as follows:

(1) Will a Picture be Worth One Thousand Words on Telidon?

Chapter I begins by examining a popular argument, sometimes heard in conjunction with Telidon, which says that pictures are a better, more "natural" medium of communication than words. A consideration of this issue leads to a more basic question: Are there fundamental differences between how words and images communicate? We examine this question by asking whether or not a picture can ever substitute for a verbal statement. In addition, some basic theoretical ideas are introduced which stress the flexible nature of the visual interpretation process. These theories are intended to help us better understand why pictures present difficult problems in some applied areas of research pertinent to Telidon: namely, the automatic description and classification of images for retrieval by computer, and the development of "iconic indexes."

The chapter ends by reviewing research which emphasizes that the real issue for Telidon is not how to substitute pictures for words, but rather: How the meaning of words can depend on a pictorial context, how the meaning of a picture can facilitate the comprehension of text, and finally, how visual meaning and verbal meaning can interact to generate new knowledge.

(2) Reading Pictures: Are Realistic Pictures always Best?

Chapter II begins by asking whether one must learn to "read" a picture in ways similar to learning to read written text: Can anyone who can see the visual world, see the world through pictures? Or is learning to understand pictures like acquiring a symbolic code -- a process dependent on context, culture and convention? Next, we explore the implications, for Telidon, of theoretical and empirical research which suggests that pictures departing from photographic realism -- simple outline-pictures, cartoons and caricatures -- are more easily and quickly identified than realistic photographs. We examine the notion that cartoons and caricatures work so well because they mirror the "schematic" nature of our internal concepts. Several hypotheses are put forward to suggest how this finding might enable one to exploit Telidon's strong suit as a pictorial medium: its ability to generate high quality, schematic line-graphics. In particular, we look at some possible tradeoffs, in terms of user satisfaction and ease of comprehension, among the variables of transmission time, degree of pictorial realism, beliefs in pictorial objectivity, and the spatial configuration of information.

(3) Images for Thinking: Graphic Schemes for Complex Ideas.

In this final chapter we explore some issues related to the use of simple graphic imagery to communicate complex ideas. First, we examine some of the psychological processes involved in learning and using maps -- a type of image likely to play an important role on Telidon. We point out that there may be individual differences in people's ability to use maps and find that highly schematic maps which depart from reality, may be easier to understand because they visually clarify functional connections among geographic entities. Moreover, we look at a recent attempt to replace the ordinary paper map with a dynamic "movie map" made possible by the marriage of computers and optical videodiscs. Next, we ask why it is that simple graphic schemes (Venn diagrams, boxes, arrows, etc.) are so popular in scientific and educational contexts. Their popularity, it is proposed, is linked to their ability to help us grasp generic properties and underlying relations not directly observable in the surface features of physical phenomena. While static graphic schemas are useful, it is argued that computer-animated dynamic graphics -- showing transformations -- can be even more helpful to productive thinking and problem-solving. Finally, the distinction is made between "motion" and "sequence" and it is suggested that being aware of this distinction could help systems designers devise methods for displaying dynamic visual sequences while saving transmission bandwith.

The above summary then gives an outline of the basic themes covered in the report. The reader may find that, in the course of pursuing their practical import for Telidon, I stretch some cognitive theories somewhat beyond their usual limits. I make no apologies for such theory-stretching. At this early stage of research, the general lack of knowledge about the social and psychological impact of new telecommunications systems is such that behavioural researchers are badly in need of a few bold theories which can guide or at least force the clarification of issues -- even if existing theories must be stretched a bit to do so. Also, many questions are raised in this report which are not fully answered and many interesting issues concerning pictures and Telidon are missing entirely (e.g., the role of colour). You can't do everything in one report.

In sum, the goal of this report is to be more than a literature review, but to propose a point-of-view: to advance some hypotheses which will encourage (and perhaps provoke) a deeper understanding of the role of pictures on Telidon.

CHAPTER I

WILL A PICTURE BE WORTH 10^3 WORDS ON TELIDON?

Imagine an information-provider (IP) faced with the task of creating a Telidon data-base on some topic. In addition to the problem of devising a representational structure which captures the conceptual links among the items in the data base, the IP is also faced with some difficult decisions about how to best embody those concepts and logical links in some medium which the user will readily understand: Are the underlying ideas best translated into text, pictures of various kinds or some combination of text and pictures? As we have seen in the preceding section, the choice of a medium of communication (e.g., the use of a certain picture, graphic or verbal statement) may have associated costs in terms of time, effort and even computer memory. It is important, therefore, to consider some tradeoffs which may arise between the various options available to the IP for encoding information and the subsequent benefits for the user in terms of such variables as ease of comprehension, speed of comprehension, clarity, attention-getting, and so on. But in order to better understand the nature of these tradeoffs we will first have to clarify some general questions about how people understand events and ideas through pictorial and linguistic media.

1.0. A Popular Argument about Pictures and Words

Let us begin by considering a popular line of reasoning often encountered during discussions of the role of pictures and text on videotex systems. The argument can be paraphrased as follows. We live in the era of the visual image. Pictures are everywhere (advertising, movies, magazines, billboards, etc.). People, especially children, readily and easily understand pictures. Pictures are, in some sense, "natural" vehicles of communication because they directly resemble their referents; i.e., they exploit our knowledge, acquired since birth, of the perceptual "invariants" which characterize the objects and events which are found in the natural environment (e.g., Gibson, 1966, 1971; Kennedy, 1974). In contrast, spoken and written language is seen as based on conventionalized and culture-bound symbolic codes: signs bearing only arbitrary relations to the things they refer to and which, in the case of writing, are only acquired through a long process of formal learning. In general, language is seen as some how putting greater distance between the comprehender and the referent and is therefore a less immediate and natural mode of communication than pictures. While I have oversimplified this argument somewhat, it is not unreasonable to say that a similar kind of reasoning seems to underly some rather concrete proposals in recent years which stress the advantages of pictures over words in a variety of educational, scientific and social contexts. For example: the desire to create a "pure" iconic language, not dependent on linguistic symbols, which would be universally understood by speakers of different languages and from different cultural groups (for a discussion of various aspects of this problem see Modley, 1947; Mead and Modley, 1968; Kolers, 1969; Bliss, 1965); a plea for revitalizing the process of education by using pictures and graphics -- especially computer-aided ones -- instead of verbal descriptions (Huggins and Entwisle, 1974, Ch. V); the necessity of designing man-computer interfaces which foster an immediacy and intimacy of man-machine interaction by using non-linguistic symbols (Johnson, 1970; Huggins, 1971).

In light of such enthusiasm about the communicative potential of the visual image, as opposed to linguistic symbols, one would not be surprised to find a Telidon IP asking himself or herself the question: "Given what I've got to say, could I say it better in pictures than in words?" Of course, as we shall see, there is likely to be no pat answer to the question of "better". Whether or not one chooses to use a particular type of picture or graphic versus some kind of linguistic description is likely to be dictated by specific circumstances, not a recipe: this means taking into account the nature of the "it" one is trying to communicate, who the target users are, and a host of other factors.

The question of whether and in what contexts pictures are better than words (or vice versa) is going to involve a fairly complex answer. We may get further ahead in our understanding of the relative merits of pictures and words, in the beginning, not by asking which is "better", but by asking a more basic question: namely, "Given what I want to say, can I say it at all in pictures?" In other words, perhaps a good place to start is to ask if a picture or a sequence of pictures can ever substitute for a linguistic statement. Trying to answer the question of equivalence may not only reveal some important qualitative differences between the way images and words function in communication, but may also hint at what I suspect the real issue to be for Telidon: that is, how words and images cooperate to constrain the comprehension process.

1.1. Can a Picture Ever Substitute for a Verbal Statement?

E.H. Gombrich (1972) provides us with a simple example which we can adapt for present purposes to examine the claim that verbal statements can be translated into their pictorial equivalents. Consider the sentence: "The cat is on the mat." This sentence describes a simple state of affairs. It is rather easy to form a mental image of the state of affairs described by this sentence. In fact, it is only a short step from generating a "mind's eye" picture of a cat on a mat to creating an actual picture -- in this case an outline drawing to represent the meaning of the statement "The cat is on the mat." (See Fig. 1.0). But do the statement and the drawing really mean the same thing? A first reaction is to say yes. The outline drawing clearly allows an observer (at least from Western culture) to identify the familiar objects "cat," "mat" and the spatial relation "on" as being depicted by the picture. But a more careful look at the sentence reveals some interesting differences. For example, the sentence makes clear that it is "the cat" (an individual) that is being referred to, not "a cat" (a member of a class). How could the picture by itself express this distinction? The definite article "the" places important constraints on the interpretation process; it instructs the conceptual system to try and link the noun "cat" to a particular individual in one's knowledge network which one already knows about or which will be identified shortly. We have been made aware that simple words like "the" and "a" play an important role in the comprehension process by recent work in artificial intelligence whose goal is to get computers to understand natural language. In Winograd's (1972) lan-

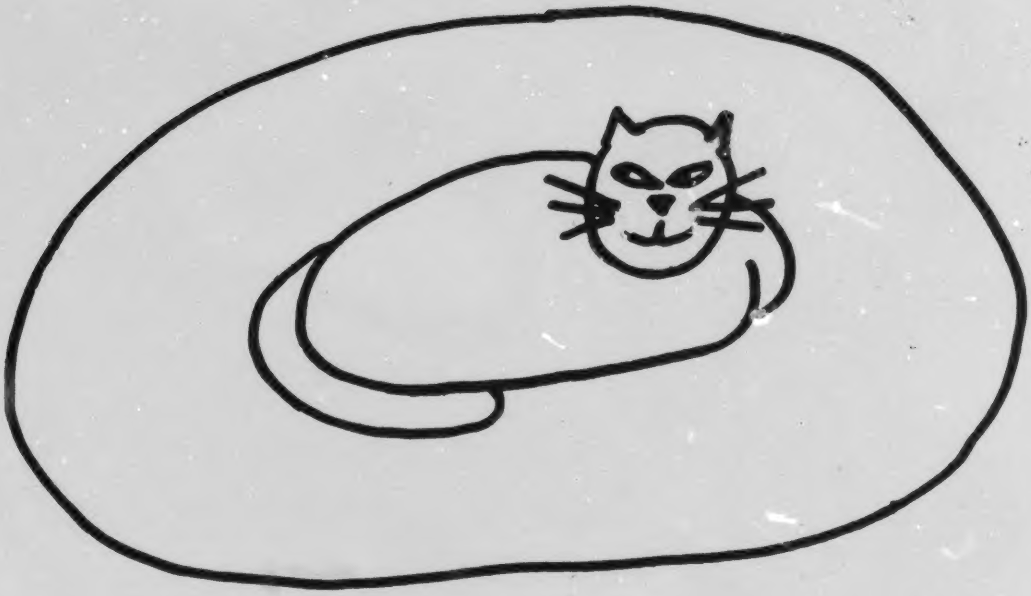


Figure 1.0. "The cat is on the mat"

guage understanding program SHRDLU, for example, the meaning of the article "the" is realized by procedures or programs which make use, not only of syntactic rules of english, but also can draw on its memory of recent events and knowledge of the world (in SHRDLU'S case knowledge about changes in the spatial relations among toy blocks) in determining whether a noun is referring to a particular individual it knows about (a secondary node in memory) or its generic definition (a primary node). Our picture, by itself, has no simple way of instructing the conceptual system as to whether the drawing refers to "the" cat or "a" cat. (This would be true even if the picture were an actual photograph of a real cat on a mat. One would still not know whether the photo was intended to refer to a particular cat or to cats in general).

If the picture in Fig. 1.0 has trouble matching the functions and "the" and "a", it really begins to fall behind the descriptive power of language when one tries to think of the pictorial equivalents of such sentences as, "The cat will get up from the mat," "There is no dog on the mat with the cat," or "If a dog enters the room, then the cat will leave the mat." Such sentences, according to Gombrich, reveal the real miracle of language: it lets us describe states of affairs, not only in the present, but in the past, future and conditional. Moreover, language contains such "logical" words as "if", "although", "but", "when" and so on, which seem integrally related to our ability to make deductions and inferences about situations. What would an image of an "if" look like (but see Arnheim, 1969, ch. 10)? Of course, one could argue that while a single picture may not be able to accomplish such feats, it may be possible to devise a sequence of pictures which could depict past, future and conditional states. For those who believe this, a useful exercise would be to try to invent a purely pictorial sequence, using no linguistic symbols, as an equivalence for the sentence "The cat will get up off the mat." We predict that such an exercise will quickly boost one's admiration for the efficiency with which verbal language performs its descriptive function.

1.1.1. Pictures Are Worth 10^3 Words, But Which Ones?

Let us try to be very clear at this point about the issue we are raising here concerning pictures and language. We are not claiming that a picture is necessarily inferior to a verbal statement. Nor are we claiming that pictures are unable to provide useful information about objects, situations and events. What is being claimed is that there does seem to be a basic problem with using a picture as a substitute for a verbal statement. The problem is not that a picture is not worth a thousand words. A picture may indeed be worth a thousand words: the real problem is that without some constraining context, we often do not know exactly which ones. In other words, for any given picture, even a simple one like Fig. 1.0, we can read off of it an endless number of possible propositions. Pictures, by themselves, have no simple way of guiding the comprehender to the subset of propositions that was intended by the sender.

To summarize: in this section we have raised the question of whether a picture could ever substitute for a verbal statement and saw that one of the problems with pictures as a "stand alone" medium is that they have no simple way of constraining the infinite number of meanings that can be read off of them. In the next section we will explore this problem of multiple meanings in pictures further by relating it to some views about how perception functions in general. These views will highlight the flexibility and context-sensitivity of the perceptual process and will help us better understand why pictures pose difficult problems in a number of applied areas which may be relevant to Telidon: for example, how to classify pictures or scenes for information retrieval by computer, or what is involved in developing "iconic indexes." These examples will reveal the importance of context in the comprehension process and, in our opinion, point to some important issues concerning the use of pictures and language on Telidon: how language serves as a context for understanding scenes; how a picture constrains the interpretation of text and finally, how pictures and words can interact during comprehension.

1.2. Multiple Meanings in Pictures: A "Schema" Approach to Perception

We have been arguing that the meaning of a picture -- even a very simple one like Figure 1.0 -- may be more or less clear; that a problem with any picture as communicative device, is that by itself, it cannot direct a comprehender to the particular meaning, among an infinite number of possible meanings, which was intended by its sender. One might ask, however, if the problem of multiple meanings is unique to pictures? Or is it characteristic of visual interpretation in general? For example, does the problem of multiple meanings arise in the everyday perception of events and objects in the visual world? Certainly, in moving about in the visual world we rarely have difficulty in identifying the objects, people and events which surround us. This is so despite the fact, as pointed out by classical theorists of perception (see Appendix A) that a given two-dimensional shape projected onto our retinas might have been produced by any one of an infinite number of differently shaped objects at varying slants and distances from the observer. Moreover, the information in the sensory array is in constant flux as we move about in the world due to changes of viewpoint, other objects becoming interposed between ourselves and a perceived object, temporary changes in illumination, and so on. One of the great challenges to perceptual theory has been to explain our ability to maintain a stable perception of the world despite these constant shifts and changes.

One approach to these problems of multiple meanings and constant flux is to deny that they are really problems at all. This is the stance taken by J.J. Gibson and his followers (for recent discussions of the Gibsonian position see Shaw and Bransford, 1977; Kennedy, 1975) who point out that our perceptual systems rarely encounter problems of ambiguity in everyday perception because the visual system is "tuned" to "pick up" important higher-order informational invariants in the optic array which specify, automatically and unambiguously, the objects, people and events in the natural environment. Instead of seeing shifts of head and body position as problems, the Gibsonians see these changes as providing valuable "time-varying" information to the visual system -- patterns of optical structure potentially describable by the laws of "ecological physics" (e.g., gradients of texture and illumination, occluding edges and corners, etc.), which resolve sources of apparent ambiguity inherent in still pictures. A gibsonian view of perception, sometimes called "direct realism," therefore underplays problems of ambiguity, experience and the role of mental inferences in perception -- elements central to what is sometimes referred to as the "cognitivist" or "constructivist" approach. A detailed discussion of the constructivist/direct-realist debate is beyond the goals of this report (but see Appendix A for a review). For present purposes, it is sufficent to say that because of their unwillingness to acknowledge the contribution of memory, learning and mental inferences in perception, we have found the gibsonian approach somewhat less useful as a framework for thinking about the problem of pictures on Telidon than a constructivist view (but see the contribution of Kennedy, 1974 in Chapter II of this report).

1.2.1. Schemas and Transformations

Without going into great detail let us briefly sketch one particular constructivist approach to perceptual understanding which may shed light on the problem of multiple meanings in pictures. This approach is based on the notion that perception, and understanding in general, involve a kind of "problem-solving" or "theory-building" in which the conceptual system attempts to account for the raw data reaching the senses by the forming of a "description" -- a kind of mental model or theory of the sensory array which explains it by fitting it to a set of underlying concepts (Bregman, 1977, 1979; Rock, 1975; Minsky, 1975). Following Bregman (1977) we can think of the mental units or concepts out of which descriptions (theories) of sense data are built as being of two basic kinds: "schemas" (Bregman calls them "ideals") and "transformations." A schema or ideal can be defined as a mental structure which encodes the brain's knowledge of any regularity or consistent pattern of experience; these structures are seen as being like "formulas" or "stereotypes" which capture our prototypical knowledge of objects, events and situations. Now while schemas or ideals capture prototypical knowledge of entities, transformations are concepts which account for the changes or modifications which may occur to the prototypical form of schemas. Transformational concepts are necessary because there will always be some mismatch between the mental ideal or prototype representing an entity in the brain and the exact way an entity may occur in any specific real-world scene or situation. In visual perception such mismatches may occur because of the particular viewing angle of the observer, the presence of other objects in the scene, changes in illumination and so on. A simple example of how schemas and transformations operate in visual interpretation would be the conceptual system's ability to account for an elliptical shape projected on the retina by interpreting it as a schema -- an ideal circle form -- which has undergone a slant transformation due to the observer's particular viewing angle. The schema plus transformation approach to perception is an attempt to show how the conceptual system could use a finite set of general concepts or frameworks in order to understand an infinite variety of specific scenes and events. This idea is not new to psychology; it can be found in the work of Bartlett (1932), Piaget (1947), and more recently Minksy's (1975) "frame theory." It is also an attempt to model Chomsky's (1965) insight in linguistics that a finite set of components of "deep" structure plus rules of transformation could account for an infinite number of novel sentences at the surface level.

Schemas and transformations also function in our interpretation of temporal events. This can be seen with reference to the cartoon-strip in Fig. 1.1. One possible low-level description of this drawing is as follows: "Starting from the left, there are four adjacent rectangles. Each rectangle contains a black dot and each dot is bigger than the one to its left. The final rectangle is black." This is one possible description of the drawing. Yet as a theory of the sensory data, it is fairly unparsimonious. Most likely our cognitive systems will prefer to account for this drawing in a more elegant and economical way by employing high level transformational concepts. For example, instead of seeing the drawing as representing four pictures of four distinct and independent objects, our knowledge of the rules governing perceptual change, combined with our knowledge of the rules of reading cartoon-strips, yields a mental representation of the cartoon in which the separate forms are seen as a single "ideal" object undergoing some type of transformation: i.e., for example seeing the changes in size and shape as representing a single object moving toward a stationary observer. Such a transformation would seem to underly interpretations of the cartoon as representing such events as "looking at a charging bull" or "getting hit on the nose by a ball." "Motion," then, in this example, is a transformational concept which, when inserted into one's mental representation of the drawing, enables the conceptual system to see a complex pattern -- the separate dots of different sizes (including the blacked out rectangle) in terms of simpler, underlying regularities -- i.e. an ideal object in motion.

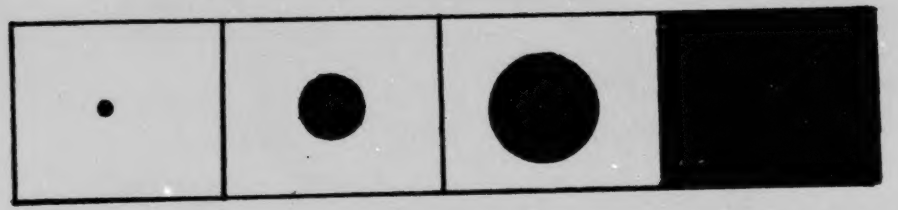


Figure 1.1. A cartoon-strip analogy

From: Mills (1980)

The schema plus transformation approach to visual understanding can help us see why we can read an endless number of different propositions off of a picture. If perception is a type of problem-solving in which the conceptual system attempts to account for or explain a visual pattern by forming a theory or description of it, built out of a finite set of underlying schema, then the same pattern or "data" can be accounted for by competing theories -- i.e., by different combinations of underlying concepts. Exactly which combination of concepts will be chosen in any given situation will be dictated by the goals and expectations of the comprehender, the surrounding context, etc. The ability to generate alternative descriptions for the same pattern or event -- especially as a function of changes in context -- gives our cognitive systems flexibility and is most likely at the heart of our creative and problem-solving skills.

One can see the flexibility of the interpretive process most clearly when contextual constraints are minimal and when the visual elements of a picture are relatively abstract. Consider, for example, some of the ways in which college students were able to interpret the cartoon in Figure 1.1 when asked what kind of familiar event the drawing reminded them of (for a full discussion of the psychological processes involved in generating verbal descriptions of "abstract" cartoon-strips see Mills, 1980a, 1980b).

1) getting a rock thrown at you
2) getting a ball on the nose
3) looking at a nail-head from different heights
4) seeing a planet grow bigger as I descend in a space-ship

5) an expanding circle caused by a pebble dropped into a still pond

6) a non-resolvable problem you have to live with

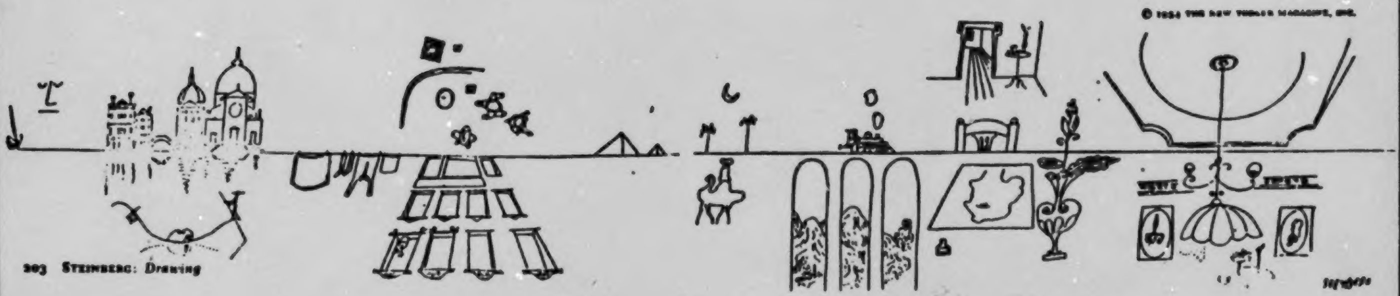
While there is no space for detailed analysis of these interpretations, it should not be difficult for the reader to see how different combinations of schemata and transformations (including assumptions about the perceiver's own role in the event) are able to account for the shapes and patterns of changes in the drawing. For example, the pattern of changes in Fig. 1.1 can be explained not only by seeing them as a single object moving toward a stationary observer (interpretation 1) but also as a moving observer approaching an immobile object -- a different kind of transformation embodied in the interpretation, "seeing a planet grow bigger as I descend in a space-ship." Other interpretations are built out of different combinations and schemas and transformations.

Now it is true that in this example the simplicity and abstractness of the cartoon as well as the lack of contextual constraints encourages a wide variety of possible descriptions -- and who can say which is best? We believe, nonetheless, that there is no fundamental difference between finding an interpretation for an "abstract" cartoon such as Fig. 1.1 and the interpretation of ordinary pictures whose representational elements attempt to depict the appearance of recognizable objects. In both cases the conceptual system must have concepts or rules (schema and transformations) which permit a mapping from the representational domain (lines, patches of color on paper) to the target domain (recognizable events, 3-D objects, etc.). And, as we shall see below, for any visual element -- whether a realistic picture or an abstract cartoon -- different mappings are possible for the same input given changes in context, including the comprehender's goals and expectations. This fact has been amusingly portrayed by the cartoonist Saul Steinberg who in a single image, shows how a simple graphic element -- the line -- can take on very different meanings by varying its context. (see Fig. 1.2)

1.3. Picture Descriptions for Machines

The implications of a "description" approach to visual perception which emphasizes the non-rigid nature and context sensitivity of the interpretive process are not purely theoretical. One area of applied research which has been forced to deal with the problems of flexibility and context in visual interpretation on a very practical level, is the attempt to develop information retrieval systems where some of the items to be retreived are pictorial in nature. (e.g., Firschein and Fischler, 1972, 1969; Fischler, 1969; Narasimhan, 1969, Clowes, 1969). This work would enable human users to query a system about the existence of a particular picture, or set of pictures, including questions about some subset of properties which the picture may contain (or may be inferred to contain): for instance, such a system might have to answer a question, based on a stored description of a picture, such as: "Does it contain a waterway?"

Firschein and Fischler (1972) have likened the problem of developing a question-answering system using pictures as data-bases to the


© 1954 THE NEW YORKER MAGAZINE, INC.
Steinberg

203 STEINBERG: Drawing

Figure 1.2. A Steinberg cartoon

Source: The New Yorker Magazine (1954)

problem of preparing an "encyclopedia" -- i.e., how to structure a large body of knowledge in such a way that it can be conveniently accessed by different types of users for different purposes. Creating such an encyclopedia based on knowledge derived from a picture requires nothing less than finding a way of representing the "meaning" of any given picture in some type of symbolic structure -- for example, a semantic network which uses a finite set of nodes and links to capture -- in canonical form -- the conceptual relations among items in the picture (e.g., x is west of y (location), x is a y (class inclusion), x stretches across y (operations), etc.).

One method of building encyclopedic representations of pictures is to ask a group of human subjects, with different backgrounds, goals, level of expertise, etc., to generate verbal descriptions of a given picture (e.g., an aerial photo of some territory). One then maps the descriptions, according to a set of rules, onto the underlying semantic relations. The semantic structures derived from the descriptions of different subjects are then combined into one "composite" representation which, theoretically, could be used to answer questions about the picture from many different viewpoints or contexts (e.g., geographical, geological, political, military intelligence, etc.). Apart from the problem of dealing with the immense number of possible contexts or viewpoints (different underlying schema), the problems encountered in creating such a "composite" encyclopedia, based on human descriptions, are immense. For example, one person may describe a portion of a picture as "a house with a swimming pool and tennis court" -- a relatively objective labelling of a group of objects in the picture. A second person, however, may describe the same objects as depicting "an expensive house" (Firschein and Fischler, 1972) -- clearly a deductive step which takes them beyond the information given explicitly in the photo. The fact that, given a picture with a finite set of objects, there is virtually no limit to the number of novel inferences (new descriptions) which could be generated, would appear to be a major obstacle to the task of creating an "encyclopedic" representation of the meaning of a picture. An analysis of the problems (and solutions) posed by the automatic description and retrieval of scenes is beyond the scope of this report and the expertise of the author. The problem area is, however, of obvious importance to the future of Telidon because one can easily imagine situations where people may want to use Telidon to retrieve pictures of various kinds as well as information about pictures without necessarily wanting to see the picture itself.

1.4. Can Pictures Serve As Indexes?

We have been examining the question of whether a picture, by itself, might be able to substitute for a verbal statement. We have been arguing that the main problem with using pictures as substitutes has its roots in the very nature of the human conceptual apparatus -- an extremely flexible device which can draw on a finite set of concepts (schemes), and compose them together in varying combinations, to generate an infinite number of different "descriptions" (mental representations) of a given visual pattern whose source is either a pictorial surface or an object in the world. Yet there are examples of pictorial communication where the conceptual system does not crank out new meanings endlessly, but where the meaning of pictures are clear, even where verbal language itself is of little help. The best example of this can be found in international events like the Olympic games where pictorial signs function as indexes of sorts; that is, they indicate to the observer, regardless of their native language, by visual means alone, which particular athletic event is being referred to. These pictorial indexes are often quite ingenious; they are often abstract forms, simple yet easily recognizable as representing people, objects and actions. We will look more closely at some of their properties shortly.

E.H. Gombrich has reminded us that the usefulness of iconic symbols such as the Olympic signs is greatly dependent on their having a clear context of use. Where the context is confused or unclear, the use of pictures as indexes or labels can have undesirable, if sometimes comic, consequences. Gombrich recounts:

> Some years ago there was a story in the papers to the effect that riots had broken out in a underdeveloped country because of rumors that human flesh was being sold in a store. The rumor was traced to food cans with a grinning boy on the label. Here it was the switch of context that caused the confusion. As a rule the picture of fruit, vegetable or meat on a food container does indicate its contents; if we do not draw the conclusion that the same applies to a picture of a human being on the container, it is because we rule out the possibility from the start. (Gombrich, 1972, p. 87).

The Olympic signs avoid such confusion because the context is clear, and thus can function successfully as indexes even where words do not help. But even where context is clear we can ask whether some kinds of pictures function better as indexes than do others, a question potentially relevant to the possible use of some form of pictorial indexing on Telidon. In order to examine this particular question, however, we must make a brief digression to consider a difficult question that we have avoided until now: namely, what functions do pictures perform in general and whether certain functions are best served by certain visual properties? Fortunately, this question has already been treated by Rudolph Arnheim (1969) and we will draw heavily on his insights.

1.4.1. Pictures, Symbols, Signs.

Arnheim makes the useful distinction between the physical properties of an image and the communicative function it might fulfill. This distinction between form and function can perhaps best be seen by showing how the same image can serve different functions -- sometimes simultaneously. Arnheim has said that an image can function as a picture, a symbol or a sign. Take a simple triangle for example. A triangle functions as a picture when it is meant to refer to a mountain; as a sign when used to indicate danger; and as a symbol when it is used to represent the idea of a hierarchy. Briefly, an image functions as a picture, according to Arnheim, when it represents things at a lower level of abstraction than it is itself. So, for example, one could use a triangle -- an abstract shape -- to depict a mountain. The "picturing function" is possible because the triangle has structural features which are partially isomorphic with those of actual mountains (base, summit). But a triangle is schematic and simplified compared to a real mountain -- although they share certain structural properties. An image behaves as symbol when it represents things at a higher level of abstraction that it is itself. For example, the structural features of a triangle could be used, not only to depict appearances, but to represent an idea: for example, to depict the logical notion of "hierarchical organization." In Arnheim's words, "symbols give particular shape to types of things or constellations of forces." An important class of images to be considered later on, which fulfill symbolic functions, are abstract or schematic graphics such as waveforms, Venn diagrams, arrows, boxes, etc. which give visual form to purely abstract patterns of forces, intensities, strengths, vectors, overlapping sets, and so on. Finally, an image acts as a sign when the link between itself and its referent is arbitrary -- merely one of association, in which the structural features of the image bear no isomorphic correspondence to the structure of the things it refers to. Words, of course, are the examples of signs, par excellence. The word "cat" does not bear any physical similarity to a cat. Sometimes the question of whether an image functions as a symbol or a sign is not so clear cut. For example, calling our triangle a sign for danger, as we did previously, is somewhat tricky. A circle might serve as effectively as a sign of danger as a triangle. All that is really required is that the circle be clearly distinguished from other shapes and used in a consistent manner. But there is something about the sharpness of triangle that makes it a better sign for danger than a circle; i.e. the abrupt turns and points of the triangle seems to evoke abstract structural qualities at a deeper level of abstraction and thereby functions both as a symbol and sign simultaneously.

We are now ready to return to the problem of pictorial indexes and whether some types of pictures fulfill this indexing function better than others. First of all, what makes for an efficient indicator? A good indicator, as we have hinted above, should be one which quickly and unambiguously selects its referent from among a large set of alternatives. Now a first reaction might be to say that highly realistic pictures would serve this selecting function best. Yet is this really so? Think of the research on automatic scene description discussed in the previous section which showed that even a realistic photo can give rise to an endless number of alternative descriptions. Part of the difficulty may be that the more realistic a picture, the greater may be the competition among its traits for the attention of the observer. Arnheim describes how this problem was encountered during an attempt to use life-like images as traffic signs. Consider using a realistic picture of a snail to signify "reduce speed." While it is true that a snail is usually thought of as a slow-moving creature, a realistic picture of a snail could give rise to other qualities as well -- slimy, small, frail, patient, etc. The highway context may help converge the observer's attention onto the correct alternative, but the picture itself does not perform this converging operation.

A second problem with using realistic pictures as indexes is related to the way our internal classification systems are structured. A given object may belong to a large number of supersets and so a picture of an object, while it may be clearly recognizable, has no way of specifying its precise superset. A picture of a wrench accompanying a set of menu items on Telidon may signal the user that there is a body of information concerning things mechanical in general, but how will the user know whether the wrench indicates "plumbers," "auto mechanics" or "hardware stores"? A picture of a wrench may be attention-grabbing, but may easily set up false expectations on the part of the user as to which superset, among the possible supersets, is intended. A serious challenge for Telidon IP's will be to evaluate this tradeoff between the possible attention-getting liveliness of a realistic picture and the probability that it wastes the user's time by activating the wrong superset expectation. It may be that, unless the context is quite clear, the word for the superset will be the most simple and efficient means of indicating a particular category -- an accompanying picture being redundant or, worse, slowing down the search through the data base by demanding the user's attention and by leading him or her down a garden path of false expectations. Of course, these are empirical questions that could be investigated experimentally.

One body of work in cognition that might prove useful in thinking about problems of indexing -- whether based on words or images -- is research into the nature of human categorization (Rosch, 1975; Rosch and Lloyd, 1978). This work has shown, for example, that the members of a class may not be equally good representatives of that class -- i.e., some items are more "prototypical" than others. Thus, if you wanted to activate the superset "furniture" by a particular example you would do better by using the terms "table" or "chair" than by using "sofa" or "bench." Presumably this would hold true when using pictures of various objects instead of words to activate the superset. An intriguing example of how this may work pictorially is recent work on categorization in American

Sign Language (ASL) -- a purely "iconic" language used by the deaf (Newport and Bellugi, 1978; Bellugi and Klima, in press). This language is based on the use of hand gestures or signals, not to spell out the letters of words, but to trace "pictures in the air." For example, the sign for "guitar" in ASL is made by strumming an imaginary guitar. These hand gestures can be used to communicate, not only about concrete objects (chair, guitar), but also about abstract ideas (democracy). Now one of the interesting properties of ASL is how, using purely iconic gestures, one can activate in the mind of the receiver the notion of a superordinate. For example, how can one represent the idea of "furniture" or "musical instruments"? What one finds is that in ASL there is no single sign for a superset term. What the signer does is string together the separate signs for several individual objects which are prototypical members of the class. The only change is that the rhythm with which the separate signs are made is accelerated. For example, to say "musical instrument" in ASL the signer would first make the gesture for "guitar" quickly followed by the signs for "piano" and "trumpet." -- i.e., the prototypes which activate the superset. There are ways of visually signalling the subset as well (i.e., upright piano) but we do not have space to treat this topic in detail.

What does this use of signs in ASL tell us, in practical terms, about using pictures as indexes on Telidon? First, it suggests that if a picture is to serve an index one should verify its "prototypicality." Second, is that two or three pictures of prototypical objects would be better than one for indicating a superset category? There may be other useful insights concerning the use of pictures and graphics on Telidon by studying the nature of ASL. In sum, the work on human categorization appears to be very relevant to the problem of indexing on Telidon -- whether using pictures or words.

While pictures may be useful as indexes on Telidon, will life-like images always be the best indexes? Not necessarily. A life-like image may be easily identifiable as depicting a particular object, but may not be the best means of quickly indicating the generic property characteristic of a category of information. To get at the generic meaning pictorially, perhaps more simplified, abstract images would be better. A diagrammatic arrow, Arnheim tell us, is a better symbol for pointing than "a realistically drawn Victorian hand with fingernails, sleeves, cuffs and buttons." Abstract images -- symbols -- are able to focus our attention on generic properties with greater precision than do realistic images. The problem with a highly abstract image, however, is that while it can highlight general properties it can cast too wide a net. For example, our triangle, while simple, abstract and easily identifiable, has structural features which can be fit to many things (be assimilated to many underlying schema) -- a picture of a mountain, the shape of a house, the notion of hierarchy. It could even be used to represent the class of triangles.

The best pictorial indexes may be those which fall somewhere in the middle of an imaginary scale of abstraction with realistic reproductions at one end and highly abstract images at the other. An example of such a mix of the abstract and pictorial are the highly succesful iconic indexes uses during the Montreal Olympics. (See Fig. 1.3).

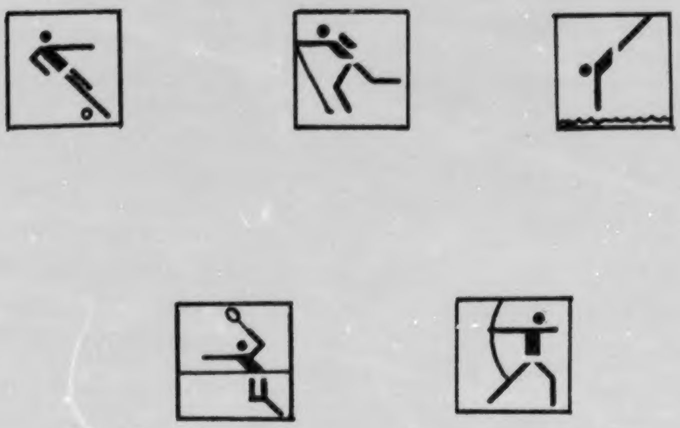


Figure 1.3. Iconic Indexes Based on Those Used During the Montreal Olympics.

These images are visually striking, easily identifiable (within certain limits) and yet do not attempt to mechanically reproduce the appearance of the entities they portray. In fact, these images are built from sets of simple geometric shapes: circles, squares, rectangles, and the like. What makes them powerful indicators is not only that the context of use is clear, but also that they perform two functions simultaneously. Not only do these designs convey the pictorial qualities of the objects associated with a given athletic event, they also seem to embody the more abstract underlying patterns of forces -- intensity, strength, movement, balance -- which form part of our mental concept of a particular sport: qualities having no measureable physical correlate in the real world. For example, the left-most picture in Figure 1.3 seems to capture the "visual dynamics" of a soccer player kicking a ball in a way that no actual photograph of a real soccer player could do. The abstractness and simplicity of the graphic elements help us see the generic forces of motion, balance, force, and so on, underlying the event -- yet at the same time manages to be a "picture" of a soccer player. Intriguingly, we may find that an efficient iconic index will possess certain properties characteristic of all great art. It will function simultaneously as a picture and as a symbol; that is, a good pictorial index may be an image of a recognizable object or event, yet convey the "deeper" pattern of generic forces underlying them.

1.5 The Importance of Context

If there has been a single message in what we have said thus far, it is that it would be a mistake to try to use pictures as stand-alone substitutes for verbal descriptions: this would be to stage a fruitless contest between words and images and only manage to draw attention away from the real-issue: the importance of context in comprehension; namely, how words can serve as a context for facilitating pictorial understanding and how pictures can aid linguistic comprehension. Cognitive psychologists in recent years have been more and more concerned with understanding the role of context in comprehension. Many experiments have been conducted which demonstrate that without appropriate context, a linguistic or pictorial message may be difficult to understand, less easily remembered, recognized, and so on. For present purposes we are mainly interested in experiments which show context effects across the symbolic modes of language and images.

1.5.1. What Words Do For Pictures

Let us look first at how language can provide a context for pictorial comprehension. In their now classic experiment, Carmichael, Hogan and Walter (1932) showed how a linguistic description can influence people's reproductions of simple outline drawings (e.g., two overlapping circles) seen in brief exposures using a tachistoscope. By simply seeing the name of one of a pair of possible objects (eyeglasses versus dumbell) before the exposure of the picture, subjects tended to reproduce a drawing which resembled the named object. More recently, Bower, Karlin and Dueck (1975) provided an interesting demonstration of how powerful language can be as a context for pictorial comprehension. In their experiment people were asked to understand and recall "droodles" (See Figure 1.4). Droodles are difficult to understand and recall because they display unlikely fragments of objects not easily identified without some clue. Subjects receiving verbal clues accompanying the droodles performed significantly better on recall tasks than did control subjects who did not receive accompanying verbal descriptions. The authors explain the better performance of the "verbal clue" subjects by pointing out that the description provides a "schema" or "frame" -- i.e., a context which allows the assimilation of the picture into a meaningful existing knowledge structure. In terms of our earlier discussion, the sentence provides a good "theory" of the picture which facilities the problem-solving nature of the interpretive process.

These experiments showing how sentences activate knowledge structures which may be necessary for pictorial comprehension are extreme cases of what is a very important function of language in everyday communication -- although still poorly understood by psychologists -- and which is likely to play an important role on Telidon; namely, the ability of a verbal description to guide the meaning we derive from visual perception. In other words, given that we have formed a mental representation of a perceptual or pictorial event, hearing or reading a verbal description may transform this initial mental representation in important ways -- i.e., to operate on it somehow in order to derive specific, and sometimes, novel meanings. Here is a short list of some of the processes a verbal description might trigger given a picture: directing the observer

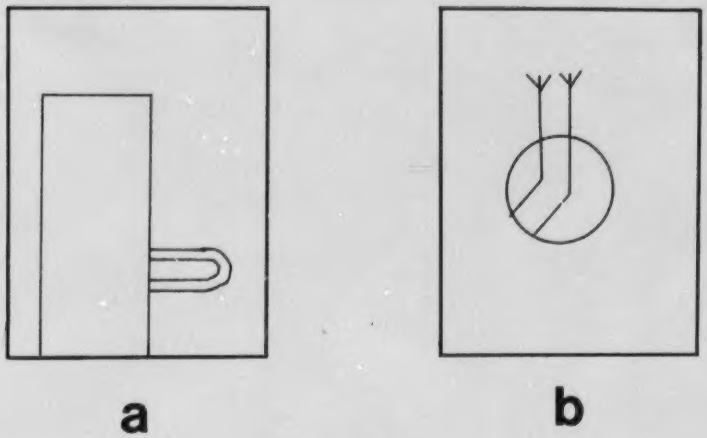


Figure 1.4. Droodles: a) A midget playing a trombone in a telephone booth; b) An early bird who caught a very strong worm.

Source: Bower, Karlin and Dueck (1975)

to select a single item in a picture among several alternatives ("look at the house with the red tile roof"); highlighting common structural features ("notice how the two bedrooms in the floor plan are both diamond shaped); pointing out spatial relations ("the water level of pitcher A is higher than that of pitcher B"); providing explanations for depicted actions ("the man in the black hat is smiling because he has just won the lottery"); helping one perceive formal structural relations in art ("the two embracing figures form a rectangularly shaped block"). The purpose of these examples is to remind us that much of everyday communication involves understanding how other people's verbal descriptions relate to the scenes we are looking at. This is likely to be the case on Telidon where instead of just having to understand a sentence or picture in isolation, the user will often be called upon to integrate information from pictorial and linguistic sources.

Sometimes the effects that words can have on the interpretation of pictures can be quite dramatic and have very practical consequences. Loftus and Palmer (1974) showed a group of people a film of an automobile accident involving a collision between two cars. After the film one group was asked to judge how fast the first car was travelling when it "hit" the second car. The second group was asked how fast the first car was travelling when it "smashed" into the second car. Some time later both groups were asked a series of questions about the accident and including, "Did you see any broken glass on the pavement after the accident? Those who had been asked the original question containing the word "smashed," said they saw broken glass more often than those who had heard the word "hit." This "leading question" experiment is a compelling demonstration of how words can activate complex knowledge structures (schemas) which serve to operate on our mental representations of perceptual events.

1.5.2. What Pictures Can Do For Words

Recent experimental work in cognition had made us aware that not only can language guide our comprehension of pictorial events, but also that a picture can serve as a context which facilitates verbal comprehension. A well-known experiment by Bransford and Johnson (1972) makes this point quite strikingly. Consider the following passage:

> If the balloons popped the sound wouldn't be able to carry since everything would be too far away from the correct floor. A closed window would also prevent the sound from carrying, since most buildings tend to be well insulated. Since the whole operation depends on the steady to flow of electricity, a break in the middle of the wire would also cause problems. Of course, the fellow could shout, but the human voice is not loud enough to carry that far. An additional problem is that the string could break on the instrument. Then there would be no accompaniment to the message. It is clear that the best situation would involve less distance. Then there would be fewer potential problems. With face to face contact, the least number of things could go wrong.

Subjects hearing this passage by itself rate it very difficult to comprehend and in fact perform poorly in tests of their recall of the main ideas in the passage. Now look at the drawing in Figure 1.5. Subjects shown this picture before hearing the passage, subsequently rate the passage as easy to understand and perform better on recall tests of the passage. The difference between the two groups, of course, is accounted for by the absence of the picture for the "no context" group; i.e., the image contains objects and relations which constitute a "conceptual

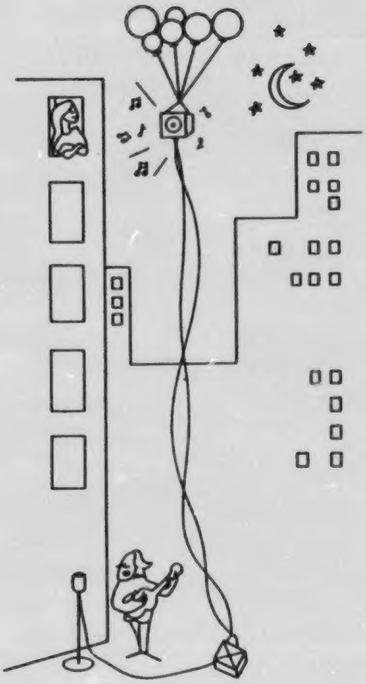


Figure 1.5. Picture from the Bransford and Johnson (1972) experiment.

base" without which the ideas underlying the sentences are difficult to assimilate. This experiment is analoguous to the Bower et al. study mentioned in the last section on how words can provide a context for pictorial comprehension.

A second body of research which can provide some insight into what pictures do for words is that concerned with the role of imagery in verbal learning. Actually, the word "learning" is somewhat of a misnomer since the predominant research paradigm here has been to ask subjects to associate pairs of words (e.g., piano, cigar) which are normally unrelated. At a later time, when provided one member of the pair, -- e.g. piano, the subject tries to recall the other member of the pair. In other words, these "paired associate" studies are not really concerned with learning in the sense of using pictures to help someone acquire a new concept in, say, mathematics, science or even to show someone how to change a tire. Still the work on imagery in verbal association does point to some interesting facts about why some images make associations between words more memorable than other images. My comments on this research will be very brief since there exists an excellent recent review by Muter (1980) on the role of pictures and imagery in verbal learning and its implications for educational applications on videotex systems.

Consider the problem of how you would form an association between the words "piano" and "cigar" so that you could recall one word given the other as a cue at a later time. Wollen, Weber and Lowry (1972) examined the role of different kinds of pictures in helping people to form such associations between word pairs. They used four kinds of pictures, examples of which are shown in Figure 1.6: (1) noninteracting, nonbizarre; (2) noninteracting, bizarre; (3) interacting, nonbizarre and (4) interacting, bizarre. They found that paired associate recall was better for the "interacting" pictures (bizarreness did not matter) than for the noninteracting pictures. What is interesting about this result, of course, depends on what it means to say that an image is "interacting." Here, unfortunately, the authors provide little help. Interpreting this result in terms of our earlier discussion, we might say that an "interacting" picture supplies the subject with a "conceptual base" or existing knowledge structure which allows them to establish conceptual relations between items which are normally considered to be separate. The study shows that pictures are effective in showing such relations and therefore aid in the association task.

1.5.3. Words and Images Interact

I have described some of the ways a picture might facilitate the comprehension of a text and, conversely, how the meaning of a text can be a prerequisite for understanding the contents of a picture. So far, though, these facilitation effects have been one-way: from text to picture or from picture to text. In closing this section I would like to point out another possible, but as yet little understood, relationship between word meaning and picture meaning: that of interaction.

There may be instances in everyday communication (and therefore on Telidon) where people are called upon to integrate information from visual scenes and verbal descriptions in complex ways. It is possible, for example, that in seeing a picture with accompanying text on Telidon,

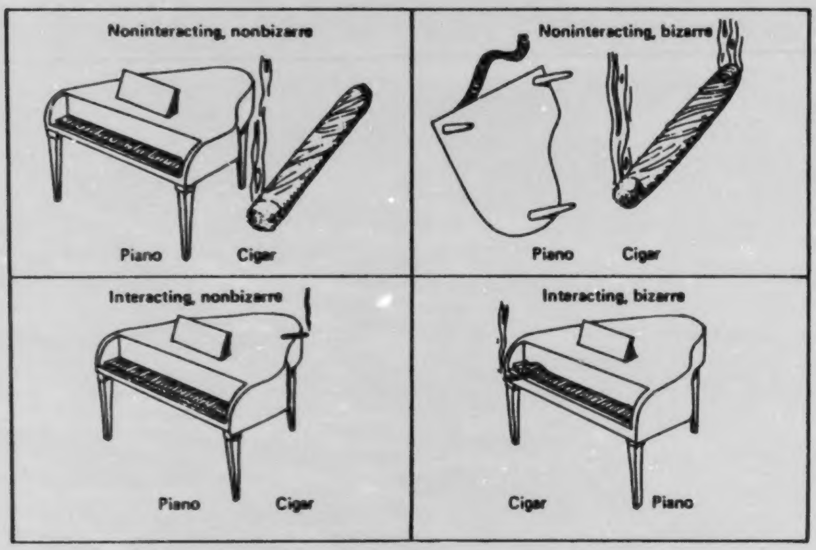

Noninteracting, nonbizarre
Piano
Cigar
Noninteracting, bizarre
Piano
Cigar
Interacting, nonbizarre
Piano
Cigar
Interacting, bizarre
Cigar
Piano

Figure 1.6. Examples of the four kinds of drawings that accompanied the word pairs presented in the Wollen et al. experiment. (After Wollen, Weber & Lowry, 1972)

that one can understand the meanings of both the text and the picture taken separately, but cannot immediately understand how they relate to each other. Such situations may arise especially in scientific or educational contexts where applying words to images is supposed to generate new knowledge. In such cases, the comprehender might have to engage in a kind of "problem-solving" involving the modification or reshaping of the underlying meanings of the words and images, in order to see how they apply to each other in the particular context.

To get a clear idea of how words and pictures can shape or constrain each other's meaning in specific contexts, let us look at an example from a recent study by Mills and Bregman (1980). First, please consider the cartoon-strip in Figure 1.7. Most people see this drawing as representing an event such as, "a foot stepping on a bug," where the black dot is perceived as being "crushed" by the downward movement of the triangle. But let's say you were asked to see how the cartoon could be an analogy for the event described by a different phrase, "a car colliding with a pole." At first, there seems to be little connection between the meaning phrase and the type of event represented by cartoon. After a brief interval, however, most people see how the cartoon could fit the event description "a car colliding with pole." What has changed to make the match possible? Usually, what changes is the observer's implicit assumptions about "viewpoint." If one imagines the cartoon as representing a "top" or "bird's eye" view -- instead of a side view as in "a foot stepping on a bug" -- then the visual elements fit the meaning of the verbal description. What is important to note, however, is that

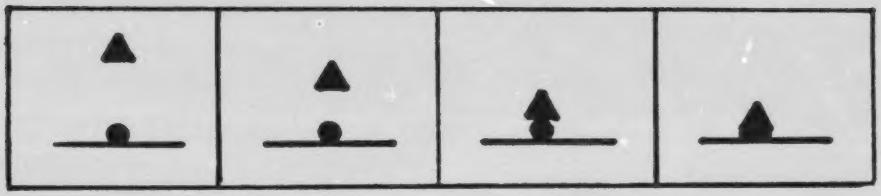


Figure 1.7. A single cartoon representing two different events: (a) A foot stepping on a bug (side-view); and (b) A car colliding with a pole (top-view).

Source: Mills and Bregman (1980)

the change to a "top-view" -- while allowing a match between the picture and the description -- brings with it changes to the meanings initially assigned to both the words and the images. Take, for example, the black dot in the cartoon. Initially, the meaning assigned to this graphic element might be limited to something like a "black, round, object" where no viewpoint at all is specified or where it is seen from a side-view. In order to fit the meaning of the word "pole," however this meaning assigned to the black dot in our mental model of the event, must be modified or updated to something like, "a black dot representing a tall, upright, cylindrical object seen from a top-view." But note that it is not only the meaning assigned graphic elements in the picture which get modified. The typical meaning of the word "pole" must also undergo some changes in the context of the picture. Specifically, if we assume that part of the prototypical meaning of the concept "pole" encodes our knowledge of the prototypical appearance of poles (e.g., tall, upright, cylindrically-shaped seen from a side-view), then some of these assumptions must be changed in order to fit the visual structure of the black dot in the cartoon. The word "pole," in this specific context, must refer to a pole as seen from a top-view in order to be consistent with the shape of the dot in the picture. The important point here is to recognize how the initial, kernel meanings we assign to pictures and words can interact or constrain each other (e.g., undergo transformations) in order to meet the demands of the specific interpretative task.

Because psychologists and linguists have tended to concentrate their attention on studying either "language" or "perception" as separate subsystems of the mind, very little is known about the kinds of interactions, such as in the example given above, that can take place between the meanings of words and images. This lack of knowledge is unfortunate because, we suspect, that such interactions are at the heart of many forms of knowledge acquisition in the arts, sciences and possibly in everyday life. In sum, the purpose of this section has been to point out that the relationships between text and images can be quite complex. There may be situations on Telidon where a picture is not used merely to illustrate a text; or where a text does not merely describe what is visible in a picture in any simple way. There a may be situations -- especially in educational settings, -- where the user is called upon to engage in a kind of "problem-solving" in order to see how a verbal description applies to a pictorial event.

1.6. Summary and Conclusion

Telidon is able to present information to the user in the form of both text and various kinds of pictures. This possibility inevitably leads to such questions as: "What is the role of visual imagery in the design of pages of information for Telidon?" "Given a choice between a textual description and a picture to communicate an idea, which one is best?" "Can I say it better in pictures than in words?" These kinds of questions, while having practical implications for Telidon, cut to the core of some complex theoretical issues in human perception and cognition concerning the nature of language, perception and pictorial representation. Unfortunately, theories of human symbolic communication are not sufficiently advanced to provide pat generalizations to questions such as "Which is better?". Questions of "better" require taking into account specific user characteristics, the nature of the task, and so on. However, our selective review of theories of cognition and perception enabled us to clarify a number of general questions concerning the nature of images and the relations between images and words which can be helpful in thinking about the role of linguistic and pictorial media on Telidon. Let us summarize these as follows:

(1) One should be wary of the popular argument that the visual image sould replace language as the primary mode of communication on Telidon -- at least as far as a "stand alone" medium. Natural language is the medium par excellence for forming propositions intended for other people. A picture, too, can be thought of as conveying "propositions." The problem is that, without some kind of constraining context (usually a verbal description) we do not know which, of an infinite number of propositions that can be read off a picture, was intended. Such semantic ambiguities make a "pure" pictorial language an unlikely prospect, in my opinion. Moreover, pictorial codes would also have a very hard time duplicating language's ability to let us form propositions about past, future and conditional events, as well as to form logical chains of inference. Again, the argument is not that pictures are not useful or even better than words in some cases, but rather that they cannot be expected to do the job alone.

(2) We found that the problem of multiple meanings in pictures had its roots in the very nature of the human conceptual apparatus -- a flexible device which uses a finite set of general cognitive frameworks or "schemas" to form definite "descriptions" of specific inputs. In visual perception, we saw that the same visual input could give rise to many different descriptions by drawing on different combinations of schemas -- often depending on changes in context. This characteristic makes the human conceptual system very supple but also makes it easier to understand why pictures cause problems in certain applied areas pertinent to Telidon such as the attempt to devise classification schemes for the automatic retrieval of pictorial information by computer; this same property of flexibility could create problems for the use of "iconic indexes" if they were badly designed. We explored some characteristics of what would make a good iconic index and proposed that a good iconic index would function simultaneously as "symbol" and "picture" and therefore share certain characteristics of all good art. That is, an efficient iconic index would use abstract graphic elements to highlight generic qualities, but would also be recognizable as a "picture" -- a visual representation of a specific kind of object, situation or event. In addition, we saw, from research in human categorization, including research into the nature of American Sign Language. that some objects are more "prototypical" members of a class than others and therefore would be more effective as

pictorial indexes.

(3) Finally, we argued that instead of trying to use pictures as "stand-alone" substitutes for words, or vice-versa, the real issue was how pictures and words function together in directing the comprehender toward the intended meaning of a message. We cited research showing that: (1) pictures can be ambiguous or meaningless unless accompanied by an appropriate verbal context; (2) a verbal passage can be ambiguous or difficult to understand unless accompanied by an appropriate picture and (3) sometimes, words and images can apply to each other in ways that change the kernel meanings of both -- especially in educational contexts. In other words, there can be interactions between word meaning and visual meaning.

CHAPTER II

READING PICTURES: IS REALISM ALWAYS BEST?

In this part of the report we will focus on what is likely to be a predominant function of imagery on Telidon, at least in its earliest stages: its "picturing" function; i.e., the use of an image to represent a thing at a lower level of abstraction than the picture itself (See section 1.4.1). By "picturing" we simply mean the use of an image to provide an observer information about what things look like, whether these things are familiar (a tree) or unfamiliar (e.g. the shape of a new tool). My goal in what follows will be to examine some ideas about the way we recognize objects through pictorial displays -- especially through outline-drawings and caricatures -- and to show how these ideas raise some provocative questions about the value of realistic, photographic images on Telidon. Specifically, I will examine how the efficiency with which the human mind understands outline drawings and caricatures may be exploited to optimize Telidon's strength as a pictorial medium; i.e., its ability to generate simple line drawings and schematic graphics.

2.0. Do We Learn to Read Pictures?

While a picture may not, by itself, be able to replace the need for verbal descriptions, a picture can provide useful information. And perhaps the most basic kind of information that a picture can provide concerns how things look: information about appearances, shape, texture, spatial relations, etc. At first glance there seems nothing mysterious or remarkable in this. A picture can stand for a real object in the world, say an apple, because in some sense, it resembles a real apple. A traditional way of thinking of this resembling process is to think of a picture as an artificially treated surface which provides the eye with the same sheaf of lightrays as would be produced if an observer were looking directly at a scene itself. In fact, this was the explanation given by J.J. Gibson (1954) of how a picture works -- by being a pigmented surface which reflects light in such a way that it can be a surrogate for real objects in the natural environment.

Gibson's explanation of how a picture works is straightforward and explains why we don't feel as if we have to learn to read pictures as we have to learn to read printed language. Anyone who has learned to see, can see a picture. But despite this common sense definition, the "resemblance" theory of picturing has had a number of critics of whom the most vocal has been the philosopher Nelson Goodman (1968; 1971). Goodman argues that the notion of resemblance does not help explain our ability to understand pictures, a process which he feels is more like learning to understand a language in its dependence on custom, convention and culture. This debate about the nature of a picture is, in a very broad sense, pertinent to Telidon because it raises the issue of the universality of pictures as a mode of communication. Will a picture or graphic on Telidon be immediately and universally understood by anyone who can see? Or is learning to read a picture a culture-bound process involving a period of adaptation or specialized training? Of course, specialized training is obviously necessary for reading some kinds of pictures: X-rays, topological maps, circuit diagrams, and so on. We will look at some of these in Chapter III. In this section we are concerned with reading realistic pictures: How we are able to identify marks on paper as a picture of a "table," "sweater," and so on.

Goodman and others (e.g., Gombrich, 1960) have pointed out that there are important differences between direct and pictorial perception. For example, two important sources of information not available in a single picture are retinal disparity (the impression of depth due to the fact that when looking at a scene our two eyes receive partially overlapping projections) and motion parallax (when we move our heads while looking at a real scene, the relative positions of the projections of objects on the retina change relative to the observer, new parts being revealed, others being hidden). Obviously, neither retinal disparity nor motion parallax can occur when viewing a picture or a painting. When we walk past a painting in a gallery, for example, the objects in a picture maintain their relative positions even as we change our viewing angle with respect to the picture. Another argument against the claim that picture perception is like direct perception concerns the perception of projective accuracy in pictures. It has been pointed out, for example, that observers are not bothered by drawings of solid objects which do not strictly follow the rules of linear perspective i.e., a cube whose titled surface has parallel, not converging lines, or an off-center disc which is circular instead of an ellipse as it would be in the retinal image (See Figure 2.0). Arguments such as these have been used by Goodman and others to support the view that reading a picture cannot be like direct perception but rather that understanding pictures must depend on custom,

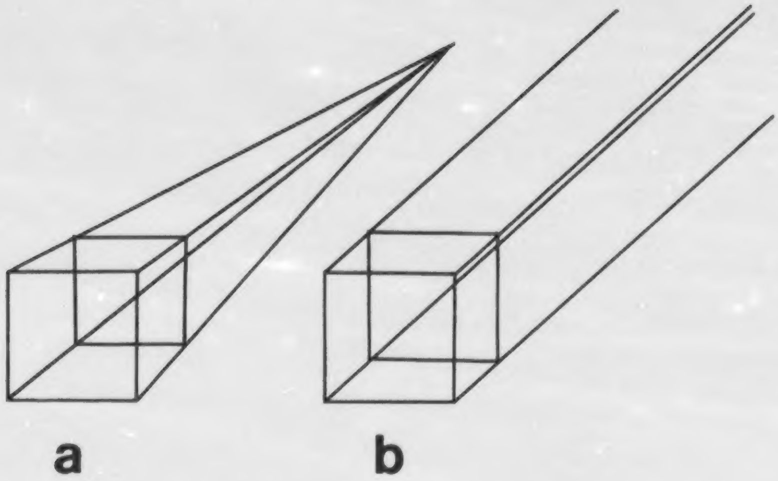


Figure 2.0. Drawing (a) can be a picture of a cube since the converging lines obey the law of linear perspective; Drawing (b) cannot, strictly speaking, be a picture of a cube since its receding edges are parallel.

convention and past experience: that pictorial comprehension is more like learning to understand the symbols of language. In a recent review of psychological research on this issue, Perkins (1979) concluded that the truth about the nature of pictorial representation most likely lies somewhere between these two extremes. Evidence can be found to support either side and usually comes from experiments with young children or cross-cultural studies of native populations unfamiliar with picturing.

One of the best-known studies using children, which supports the argument that we do not have to undergo a separate process of learning how to read pictures, is that of Hochberg and Brooks (1963). In an amazing demonstration of devotion to science, Hochberg and his wife raised their son to the age of nineteen months without letting him, as far as possible, come in contact with pictures of any kind. No magazines or picture books were allowed in the house and labels with pictures were removed from canned goods. At the end of two years, Hochberg tested his child by showing him a series of simple line-drawings, elaborate line-drawings and photos of common objects for which the child already had learned the names. The child had little difficulty in identifying these objects in any of the different pictorial modes. Support for the opposite view, that children have to learn to read pictures, comes from the work of Bower (1964, 1966) who showed that infants, conditioned to respond to the presence of real objects, would not transfer the response to pictures of these objects.

Cross-cultural studies of picture comprehension also shows mixed results, but, in general, support what Perkins has called the "real thing" theory of pictorial comprehension. That is, people from cultures not familiar with pictures, have little difficulty in identifying objects and most kinds of spatial relations in photographs or line drawings, often after a brief period of "catching on" (See Hagen and Jones, 1978, Jahoda and McGurk, 1974). An interesting exception to this, cited by Perkins, is an experiment by Duncan, Gourlay and Hudson (1973) concerning the understanding of graphic devices for indicating motion: e.g., multiple heads on the same page at different positions and streaks indicating a path-of-movement, were rarely seen as depicting motion by rural africans compared to white South Africans.

One area where there appears to be consistent differences between perceiving a picture and perceiving the real thing concerns people's ability to make judgements of spatial properties with metric accuracy: for example, Olson, Pearl, Mayfield and Millar (1976) looked at adults and children's ability to judge the slants of real boxes versus pictures of boxes. They found that both adults and children make fairly accurate judgements of orientation of real boxes but make errors in judging pictured boxes such that they greatly underestimate the real tilt by as much as 20 degrees. Perkins explains such "regression toward the frontal plane" in viewing pictures as a compromise which the visual system makes in order to reconcile the declared flatness of the pictorial surface and the depth information available in the picture. If one minimizes the flatness of the display, more accurate judgments result (Attneave, 1972).

In light of these findings then, can we make any generalizations about how pictures work? Are they a "language" bounded by custom and convention or do they work by providing information to the eye as does the "real-thing?" Following Perkins (1979) we can summarize the situation this way. Basically, the evidence supports a "real-thing" theory of pictorial comprehension, but with important qualifications. Pictures work because they package information much in the same way as do real scenes. Further, people of different ages and cultural groups can be expected to understand realistic pictures without a long period of formal learning as in learning to read a language. However, the fact that some ways of depicting require learning (e.g., motion lines) and the fact that people consistently make certain kinds of errors (in judging slants for example) means that there are aspects of pictorial comprehension in which custom, convention play a role. We should emphasize though, with Perkins, that one should not think of the role of "convention" in reading pictures as being equated with "arbitrary." For while the relation between the word "apple" and the real thing may have been decided by custom (the word "orange" might have served just as well), using streaked lines to represent motion-paths are not arbitrary in the same sense. (For example, a line of dots would not serve equally well to depict motion). Instead, the path-of-motion lines somehow capture, in a graphic device, a structural equivalent of a certain kind of perceptual invariant: the trail of lines which occur when an object passes rapidly before our eyes.

Perkins has summed up the situation this way: A "real thing" or "resemblance" theory of picturing is a good first-order approximation of how people understand pictures; a second-order theory will have to take into account the influence of custom, convention and learning.

2.1 The "Language" of Lines

A picture then is a remarkable, universally understood, device for providing people information about what things look like. (Although the particular meaning of a picture still depends on context. See Chapter I). A picture is a treated surface which packages information about the visual world in ways that are roughly similar to the way the visual world packages this information. But to provide useful information about appearances, an image does not have to be a mechanically faithful recording of a scene as in a photograph or realistic painting. One of the most powerful methods of making pictures, used since ancient times and across different cultures, which does not register each point of light in the visual array, is the simple line drawing.

Please look at the drawing in Figure 2.1. John Kennedy (1974), a former student of Gibson's has intensively studied the nature of line drawings. He has used this picture to demonstrate that many of the basic elements of optic structure can be captured by line segments. For example:

(1) An occluding bound (a curved, convex surface, part of which is hidden from the observer) with a background surface -- the brow of a hill, with the surface of another hill behind.

(2) An occluding bound with no background surface -- the brow of a hill with the sky visible above the hill.

(3) An occluding edge with no background surface -- the apex of the roof of the house, with the sky visible above the house.

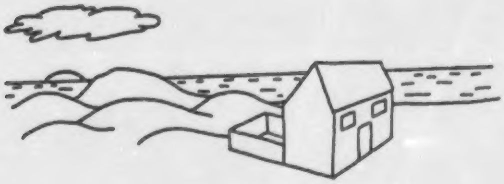


Figure 2.1. The "meaning" of a Line.

Source: Kennedy (1974)

(4) An occluding edge with background surface -- the termination of a wall, where the continuation of the surface of the wall is occluded by the rear surface, with ground surface visible beyond the termination of the wall.

(5) A dihedral angle forming a concave corner -- two visible plane surfaces of wall meeting at an angle of less than 180 degrees, measured through the air enclosed by the wall.

(6) A dihedral angle forming a convex corner -- two surfaces of a house meeting at an angle forming more than 180 degrees, measured through the air around the surfaces.

In addition to these features of surface layout, lines can depict cracks, shadows, texture, and highlights. The only features of the visual world not portrayable in lines are isolated colors and uniform space.

How can a simple ribbon of line depict so many features of the visual environment? After all, there are no lines in the natural environment, only edges, surfaces, and the like. Doesn't the ability of line drawings to portray these things mean that we must learn to read them via experience and convention? Kennedy would be quick to deny this pointing out that lines are useful because they capture, in a frozen visual array, the invariants of optical structure -- especially brightness discontinuities caused by light reflected from edges, changes in texture, surface layout, and so on. These invariants are what Gibson holds to be critically important to perceiving the visual world.

While Kennedy concentrates on an analysis of what kinds of discontinuities lines can depict, Julian Hochberg (1972) has added an intriguing twist to the story of how lines work. Hochberg begins with the insight that in order to understand the nature of line-drawings we must first take a more global functional view of perception as being a goaldirected, purposive activity which occurs over time. The key to understanding this process for Hochberg is to understand how the visual system can piece together a whole scene from successive discrete glances.

> Because our eyes register fine detail only within a very small foveal region of the visual field, we must learn about the visual world by a succession of glances in different directions. Such glances are made by saccadic eye movements, whose end-points are decided before the movement is initiated (i.e., saccades are ballistic movements): where one looks is decided in advance. Therefore, the content of each glance is always, in a sense, an answer to a question about what will be seen if some specific part of the peripherally-viewed scene is brought to the fovea. In viewing a normal world, the subject has two sources of expectations: (i) he has learned something about what shapes he should expect to meet with, in the world, and about their regularities; and (ii) the wide periphery of the retina, which is low in acuity and therefore in the detail that it can pick up, nevertheless provides an intimation of what will meet his glance when the observer moves his eyes to some region of the visual field. (Hochberg, 1972, p. 65).

Eye movements, then, according to Hochberg, are guided by expectancies or hypotheses, based on information available in peripheral vision, of what the fovea is likely to find. And what kinds of information would be most important to mobile, active organisms? The kinds of optical information described so well by Gibson and Kennedy: edges, bounds, texture gradients, cracks, etc. -- visual patterns alerting the organism to abrupt changes in the layout of the environment. That peripheral vision is sensitive to these kinds of discontinuities, usually indicated by brightness differences in the visual array, makes good sense for the survival of mobile animals. But why should lines on paper be able to elicit such hypotheses about important discontinuities. After all, as we said before, there are no lines in the world, only edges, surfaces, and so on. According to Hochberg, outline drawings probably work because the same structures in the visual nervous system that respond to the contour between regions of different brightness are those that respond to lines on paper as well. In short, lines on paper indicate the same kind of important discontinuities in the visual environment that peripheral vision is "tuned to," so to speak, in ordinary perception; the visual system is sensitive to these discontinuities because they help the organism know where to look next.

2.2 The Time/Realism Tradeoff: Implications of Using Outline-Pictures on Telidon

Imagine that you are house-hunting but that you have only a rough idea of the kind of house you are looking for, how much money you wish to spend, where you wish the house to be located and so on. Your first task is to get an idea of what is available in order to narrow down your choices. Ideally, if you had an unlimited amount of time, the best thing would be to visit, in person, each of a large number of properties that you might be interested in. But if this is not possible, a convenient alternative might be to use a videotex retrieval system as such as Telidon as a way of doing a "breadth first" search of the possibilities so as to narrow down the field to a few properties that you will eventually visit. (An ideal system might facilitate this search by using key words to provide a list of candidate properties). Now in retrieving information about the various properties using such a videotex system, pictures may have a useful role to play. Let us consider the most obvious

-- using a picture to let the prospective buyer know something about what the house looks like.

From an information provider's viewpoint, a first reaction might be to use actual color photographs of houses with as fine a resolution as the system would allow. A good photo of a house would undoubtedly be of interest to the user. But what if, in order to see this photo, on his or her TV screen, the user had to wait a certain period of time, say on the order of two or three minutes, while the photo was being built up, line-by-line on their video screen. Such delays, in fact, would occur in building up photographic images on existing videotex systems if they use narrow band telephone channels to transmit photographic quality images to the user's terminal. Each dot in the image is stored in digital code in the service center's computer, transmitted over phone lines, and mapped onto the user's terminal in a "slow scan" or facsimile mode. Thus, although one can transmit photographic images in such a "bit map" mode on videotex, one pays a penalty in terms of increased delays and storage requirements in the central computer. (Of course, one could decide to restrict the photo to a small portion of the user's screen. It would thus take less time to transmit, but one would lose visual detail).

An interesting question raised by the "bit map" method of transmitting photos is: What effect would such time delays have on the perceived usefulness of the system? If a user only had two of three houses to look at, waiting three minutes for a picture may not be a problem. But, what if, as is likely to be the case, the user wants to retrieve information on 15 or 20 houses in order to get a rough idea of what is available? Here, delays of three minutes for a picture of each house may have a cumulative effect which is intolerable. Even lesser delays of half a minute may be unacceptable in such a context, although the range of delay tolerances is unknown at this time. (Let us make clear that while our example concerns real-estate, the problems raised by the encoding and slow-scan transmission of photos generalized to other applications of pictures as well).

One can see then that while videotex systems permit the retrieval of realistic photographic images, there will be costs to the user in terms of transmission time and penalties to the IP in terms of storage requirements in the central computer. In light of these costs, our discussion in the previous section of the usefulness of outline pictures as informative displays becomes most pertinent: especially when one takes into account Telidon's ability to generate good quality line graphics. Let us explore some possible tradeoffs between time and realism.

Imagine that instead of storing the information necessary to transmit full color, photographic images of 20 houses, we were able to make contour drawings of actual houses or of particular models of houses (see Figure 2.2) and store only the "line" information instead. Significant savings could be had in terms of storage requirements for the set of 20 houses. But more importantly, since only significant features are stored in terms of simple line segments, the time it would take for a picture -- in this case an outline picture -- to appear on the user's screen would be radically reduced -- from minutes, in the case of the photo, to seconds in the case of the line drawings. Thus, a great savings in time could accrue to the user, perhaps boosting his or her satisfaction

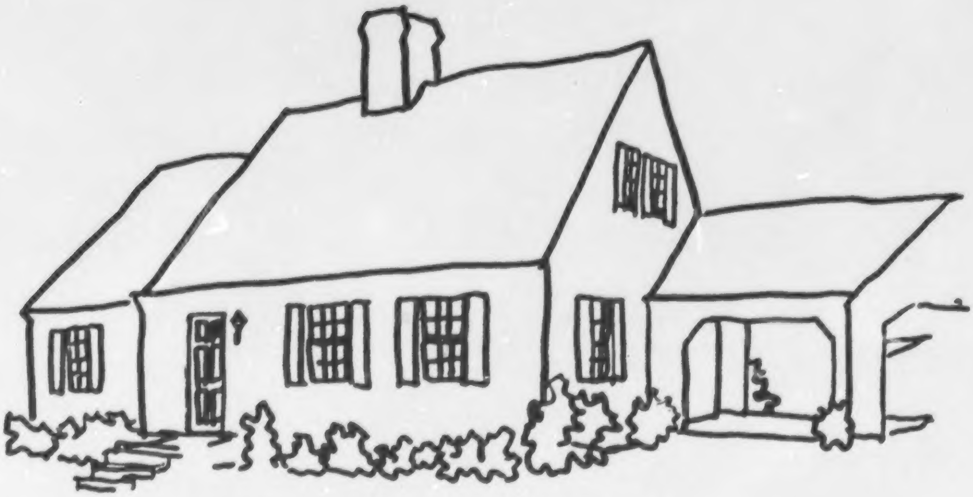


Figure 2.2. An outline drawing of a house traced from a color photograph.

with the system.[1] The important psychological issues raised here can be broken down as follows:

(1) Can an outline drawing, say of a house, serve as an equivalent representation (or perhaps be even better) than a detailed, color photo of a given scene? The work of Kennedy and Hochberg would indicate that it might; line drawings can provide information about most of the important sources of optic structure in the visual world. Indeed, if Hochberg is correct, it may even be easier and quicker to extract such information from a line drawing than a detailed photo.

(2) But there may be an important catch to this hypothesis of equivalence which relates to what might be called people's "metaknowledge" of communications media. Imagine that we have both a photograph and a line-drawing of a particular house such as that shown in Figure 2.2. Let us further assume that by some experimental test we could demonstrate that the line-drawing and the photo are informationally equivalent: i.e., that someone could answer questions about shape, size, layout, etc. equally well based upon the line-drawing as on the photo. There may still be strong

---

[1] It should be noted that "delay" has been shown to influence user acceptability of other telecommunications equipement; e.g. the time the user has to wait between the dialing of the last digit on a touch-tone telephone and the onset of the dialtone (M. Schaeffler, personal communication).

> preference for the photo for the following reason. The user may feel that a photo is a more faithful, "objective" recording of the real-thing -- not necessarily because it looks more like the real thing (you can have a poor quality, out of focus photo which is less accurate, in some sense, than a carefully done line-drawing) -- but because they believe that a "camera won't lie," although a human artist might in a sketch. Note, however, that such beliefs about the objectivity of photos might be completely unfounded. It is perfectly possible to manipulate distort, and exagggerate using a camera as any advertising photographer knows, while it is possible to use the language of line to make a relatively objective rendering of certain aspects of the information in the light.

Clearly, there are some intriguing questions here which are of importance, not only to the effective use of pictures on Telidon, but also to our general knowledge about the nature of pictorial communication: How do people's beliefs about the apparent fidelity of different kinds of pictures interact with the actual fidelity of these pictures, and how does the question of "belief" interact with such variables as transmission time, page configuration and so on.

Let me briefly sketch what some of these interactions are and how they might be amenable to empirical study. First, it would be useful to test the hypothesis that people's beliefs in the objectivity of outline versus photographic images correspond to the actual informativeness of these two modes of representation. For example, people could be presented with examples of both types of representation and asked to rate them on dimensions that would tap into their beliefs about pictorial objectivity. A second step might be to test the relative informativeness of line versus photographic representations on some simple task; for example, does seeing a line-drawing of an item, say a house, make it easier to identify that house at a later time, in another medium (e.g., photos)? A third line of inquiry might be to look directly at how a delay in transmission time affects users' evaluation of the photographic and line-drawing modes. It is possible, for example, that if a task involves using a videotex system to search through information on a large set of items, that, because of the delay factor, people would be willing to sacrifice the realism of a photo for an image which appears significantly faster on their screen, but appears in the form of a line-drawing.

In addition to type of image and transmission time, another variable which might be investigated is that of page configuration. One example here will suffice. Computer graphics technology exists today which would enable a user to view a "matrix" of separate pictures of objects simultaneously displayed on their video screens.[1] Such a matrix of images might exploit the human eye's ability to rapidly scan a spatial array to detect a particular target item from among a large set of items. Such a system would have a wide variety of applications. Again, we can use real-estate as an example. Figure 2.3 shows a matrix line-drawings of

---

[1] The Macbine Architecture Group at MIT, under the direction of Nicholas Negroponte, has been experimenting with such multiple image displays, especially looking at how the storage capacity of videodiscs can be married with the processing power of computers to provide new possibilities for picture manipulation (See Chapter III, section 3.0.1).

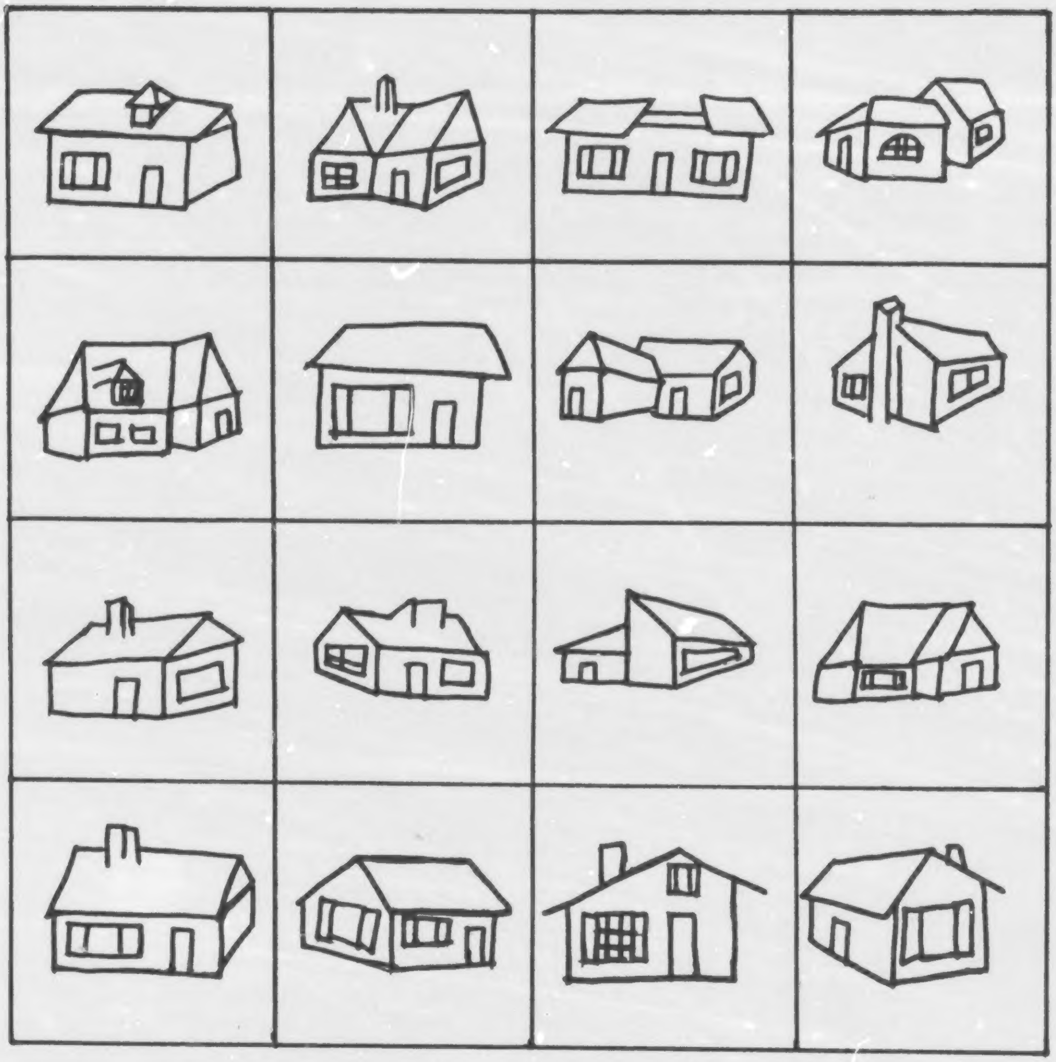


Figure 2.3. A picture matrix for "breadth-first" search.

12 houses. The medium of representation could just as easily have been photographs. A display such as this might serve as a kind of pictorial index, where a user could scan the items in the matrix, comparing, analyzing and then select one or several items for subsequent viewing in the form of larger, more detailed imagery (either photos or line-drawings) perhaps accompanied by verbal descriptions. It is not hard to see the utility of such an interactive iconic index. But should the items in such a matrix be photos or line-drawings? It should still take significantly longer to transmit a matrix of photos of houses say, than line drawings over a narrowband channel. But the interesting psychological question is, even if there were no differences in transmission time, would it be easier for the user to grasp the essential features of houses as simple line drawings than as photos. Why might this be the case? Because, as Hochberg has argued, lines highlight the essential elements (abrupt changes) which the visual system needs in order to know where to look next.

2.3. Photos, Line-Drawings, Cartoons, and Caricatures

There has been some research in education and psychology which bears directly on the question of the relative informativeness of line-drawings versus high-fidelity photographs. Dwyer (1972; 1978) has found, for instance, that the use of photographs in medical texts does not necessarily aid later identification of objects more than line-drawings -- although color was found to be useful in some cases. However, Dwyer also discovered that people expect photographs to be more informative than line-drawings, although their performance does not match their expectations.

As noted above, this question about people's beliefs about the relative efficiency of different pictorial modes deserves to be pursued in the context of Telidon.

One study by Ryan and Schwartz (1956) has directly compared the informativeness of several modes of pictorial representation: high-fidelity photographs, detailed shaded drawings, simplified line-drawings and cartoons. Figure 2.4 shows three of the four picture types used in their experiment -- the photo is missing. Subjects were presented pictures of various objects (e.g., a hand with fingers in a particular position, knife switches at various positions, views of an engine) in one of the four modes. The pictures were projected tachistoscopically for brief intervals, the exposures getting longer and longer until the subjects were able to identify certain features of the picture -- for example, the position of the fingers in the hand. Interestingly, subjects needed shortest exposures to identify the cartoon drawings; the line-drawings required the longest exposures while there were no differences between photos and shaded drawings. This study is important, not only because it raises questions about the superiority of lines over photos, but, even more , because it raises questions about the nature of cartoons? Why were cartoons the easiest to identify?

2.3.1. Cartoons Give External Form to Internal Schemes

The effectiveness of cartoons and caricatures is interesting because, from a strict Kennedy-Gibson view, they depart from a definition of what constitutes an outline picture: e.g., a delimited

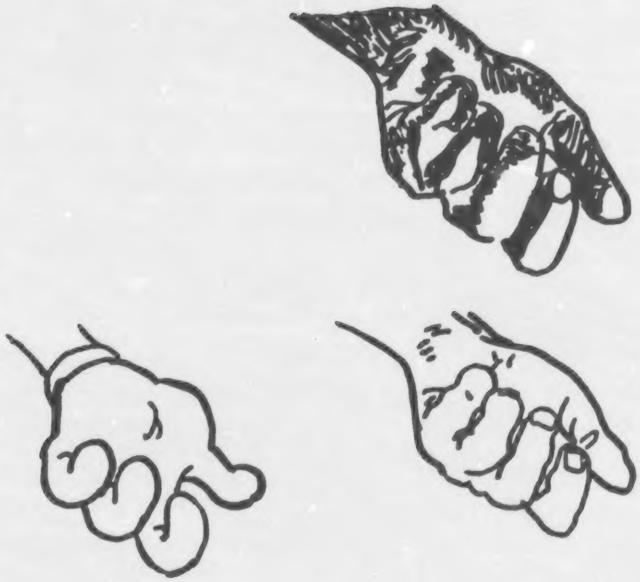


Figure 2.4. Three of the four types of pictures used in the Ryan and Schwartz experiment. The fourth picture was an actual photograph.

Source: Ryan and Schwartz (1956)

surface using lines to provide the eyes with information about a scene in the same way as would the sheaf of light rays emanating from the scene itself. Cartoons and caricatures seem to work so well precisely because they do not accurately reproduce the sheaf of light rays in a scene; they not only omit, they transform, distort and exaggerate the information in the light in ways (as yet poorly understood) that appear to facilitate recognition. The efficacy of cartoons may be somewhat less mysterious if seen in the light of some ideas introduced in Chapter I concerning the mental structures underlying perception. At that time we said that the human conceptual system encodes any input by fitting it to a combination of "ideals" or "schemas." What was implied here is that the brain interprets scenes by using "idealized," "prototypical" internal representations which never quite fit the real world, and which have to be adjusted to match any particular scene by using transformations. One theory of why cartoons and caricatures work so well is that, somehow, their external forms come closer to mirroring the internal structure of our mental "ideals" or "schemas" than do realistic photos or drawings which remain closer to the real thing. It is for this reason, perhaps, that we are fascinated and amused by cartoons and caricatures: they give us insight into the true nature of our internal schemas.

Julian Hochberg has advanced an "ideal" or "prototype" theory of cartoon perception. He begins, however, not with caricatures of faces or living things, for here difficult problems of "expression" arise, but with the problem of why some line-drawings seem to come closer than others to capturing the "essence" of the appearance of physical objects. Consider, for example, Figure 2.5 (a) and 2.5 (b). Both drawings are equally valid two-dimensional projections of a wire cube -- yet we more readily perceive 2.5 (b) as a cube than 2.5 (a). Somehow 2.5 (b) has captured more of the "essence." In Hochberg terms, 2.5 (b) comes "closer to the canonical form of a cube -- i.e., to the features by which we encode and remember cubes."[1]

Hochberg uses this notion of stored "canonical forms," together with his insights on the workings of peripheral vision, to suggest why cartoons may be more easily recognized than photos in the Ryan and Schwartz study:

> The contours (in the cartoon hand) that have been retained, for one thing, have been "simplified." That is, smooth curves have been substituted for complex and irregular ones: information about the anatomy of the hand is lost in the process (the picture is thereby made more redundant); it now requires fewer fixations to sample it and to make predictions about unsampled portions; those features that have been retained require fewer corrections to be applied to our encoded schemas (cf. Hebb, 1949; Hochberg, 1968), or canonical forms, of these objects. Moreover, the portrayed object is probably recognizable further out into the field of peripheral vision. For this reason, also, fewer fixations are required, and these are probably executed with former expectations of what will be seen.

---

[1] Hochberg's view may be seen as an alternative to the typical gestaltist approach to this problem; i.e., that the notion of a three-dimensional cube -- a simple, regular form in depth is a more parsimonious explanation of the lines in Figure 2.5 (b), and therefore preferred by the brain, to seeing the non-symmetrical arrangement of lines as a two-dimensional pattern. The arrangement of lines in Figure 2.5 (a), on the other hand, already form a symmetrical, regular pattern in two-dimensions and thus resists a three-dimensional interpretation.

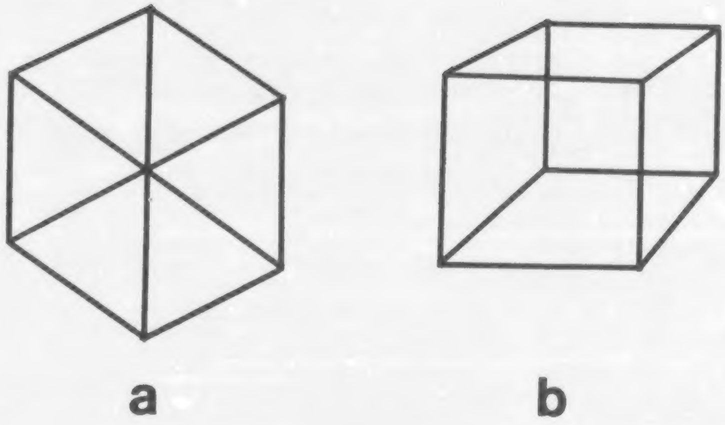


Figure 2.5. Two possible projections of a cube.

Source: Hochberg (1972)

> Although no one has actually tested the peripheral recognizability of these stimuli, the following reasons underlie the above assertions: (1) increasing the smoothness and redundancy of the curves in (the cartoon hand) must increase the efficacy of peripheral vision, which is poor on small detail and "noisy." (2) Wherever contours intersect in the cartoon drawing, the artist has deliberately rearranged them so that they meet at right angles; whether because of "good continuation," or as the depth cue of "interposition," this arrangement should make unambiguous to the observer which edge is the nearer one, perhaps even in peripheral vision. (3) Wherever the contours in the cartoon represent the boundaries or "holes" or spaces, their relative separation has been increased. This in turn has two consequences: (a) each contour is more clearly separable from its neighbors, even in peripheral vision; (b) as we have seen, the factors of contour-proximity and enclosedness tend to make a region become a figure (i.e., to be seen as an object). Increasing the separation between two fingers, for example, increases the size of the regions which we are intended to see as empty space, and thereby keeps the spaces from being seen as objects. (Hochberg, 1972; pages 75, 76)

Hochberg suggests, then, that the artist, in making a cartoon or caricature is not performing magic, nor is he or she omitting or distorting information already in the light in any simple way -- rather, the cartoonist, albeit intuitively, is performing a kind of "preprocessing" of an image, transforming it in such a way as to make it a closer fit to its canonical or ideal form in the "mind's eye." It is this closeness of fit which makes the cartoon more easily and quickly recognizable to peripheral vision.

Additional support for an "ideal" or "canonical" form theory of why cartoons may be more easily recognized than the real-thing comes from a somewhat surprising source: children's drawings. Arnheim (1964; 1969) has pointed out that, far from being "primitive," children's drawings, which typically use a vocabulary of simple forms (circle, straight line, rectangle, etc.) to build pictorial representations, actually reflect the brain's natural disposition to fit the distorted retinal projection (slants, overlaps, etc.) to canonical, idealized perceptual schemata. Arnheim would insist, for example, that while a child's drawing which employs a circle for a "head" may not look realistic, the circle captures the mind's tendency to encode the structure of the human head as being "essentially" circle-shaped. Arnheim provides many examples which show how children use simple abstract shapes to clarify and portray through the medium of drawings, the regularities and simplicities underlying an oftentimes bewildering, distorted retinal projection.

Some child artists grow into adult artists who continue to explore ways of giving external form to the inner, organizing processes of the human mind. One fascinating example is the making of caricatures. Again, what is interesting about caricatures is the non-fidelity of portrayal: i.e. they appear to break the rules of geometrical projection, in order to catch the essence of a particular individual. David Perkins (1975) has presented an analysis of the nature of caricature which helps explain the ease and rapidity with which cartoons were identified compared to photos, in the Ryan and Schwarz experiment mentioned earlier. Perkins made contour tracings of photographs and caricatures of former president Richard Nixon (see Figure 2.6). The tracing process introduces minor deviations from the original: yet the traced caricatures remain more recognizable -- are more robust in resisting the effects of such deviations -- than are the tracings of actual photographs. Why might this be so? The answer could be that, compared to the photo, the caricature does

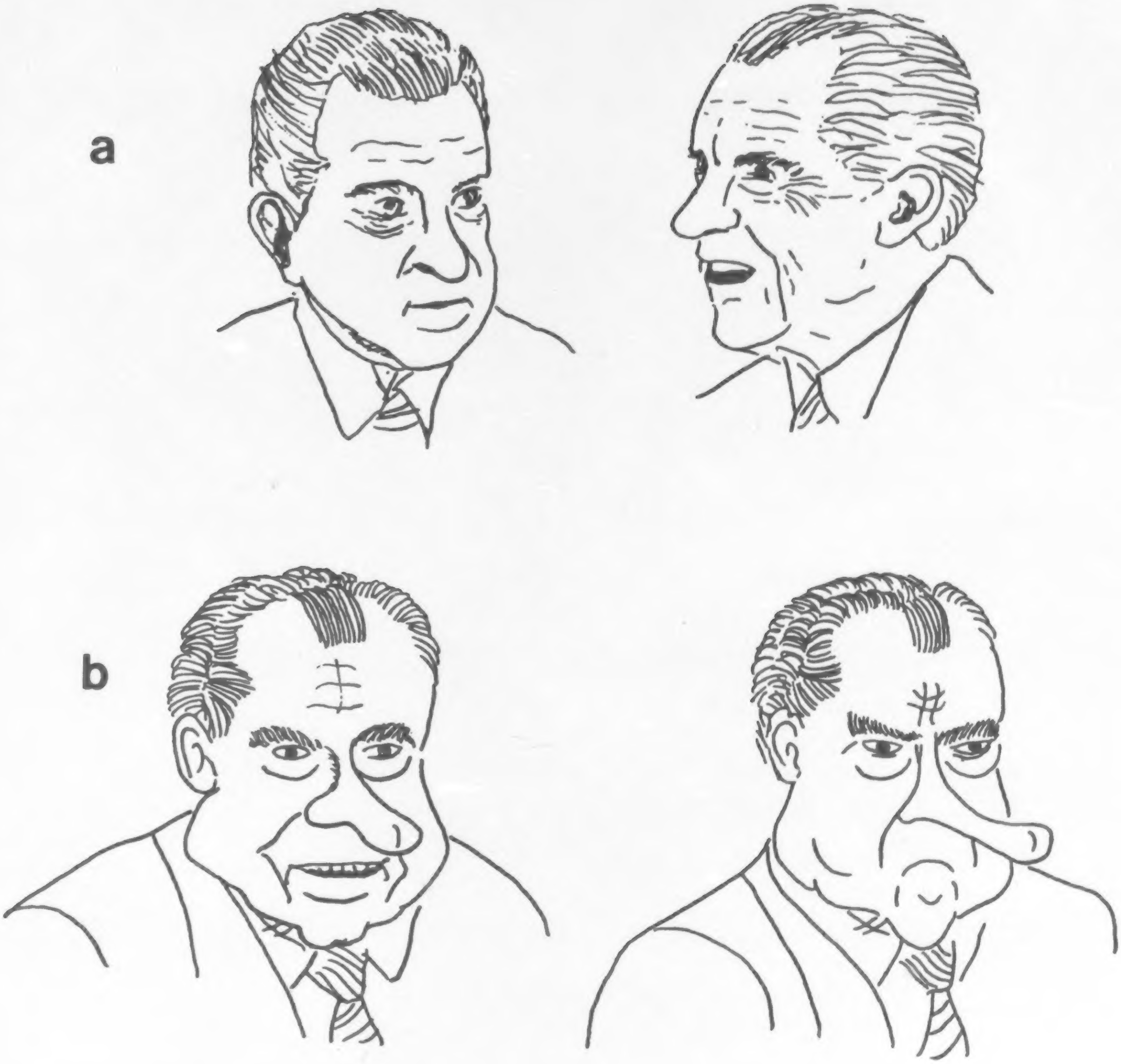


Figure 2.6. Tracings of photographs (a) and caricatures (b) of Richard Nixon.

Source: Perkins (1975)

a better job of conserving the set of "distinctive features" which we use to recognize the real face of Nixon (the jowls, the elongate nose, the hairline, the box chin) and is therefore more robust against random distortions. Again, this seems to argue that departures from realism, when they are of the right sort, can help, not hinder, recognition because such departures from real-thing is how we encode the real-thing in the first place.

There is a problem -- a kind of paradox -- with the ideal or schema theory of cartoon recognition we have been developing, however. If we recognize cartoons and caricatures more readily than photos and realistic line-drawings because they mirror the organization of our internal schema, how come we don't mistake a cartoon or caricature for a realistic drawing or portrait? Part of the solution to this paradox, suggested by Perkins, is to distinguish between the processes of recognition and resemblance.

> ...Nuances of resemblance are at issue when one already knows the identity of the depicted face, by being told or by recognizing it. Recognition involves a rapid reflexive 'look-up' system for identifying faces. Streamlined for function, the system operates with relatively few categorizations which taken together differentiate one person from another quickly and effectively, and without requiring extremely precise scanning of the stimulus. When recognition is achieved, one may then find one has available considerable further information about the face, even information contrary to the stimulus.

What seems to be implied here, is not only the difference between recognition (reflexive "look-up") and resemblance (degree of fit), but also that the mind has knowledge about its own perceptual functioning;

the conceptual system must know it uses exaggerations, simplifications and generalizations to encode knowledge of appearances, and can recognize these deviations as such when confronted with a caricature. This level of "meta-awareness" of cognitive functioning may be at the source of why we find caricatures amusing; caricatures and cartoons remind us, perhaps unconsciously, that a clever human artist has somehow tapped into the tendency of the human mind to exaggerate, generalize and simplify and has made these tendencies explicit. A similar "meta-cognitive" theory has recently been proposed in an unpublished paper by Gopnik, (1979) on the cognitive basis of humour. According to Gopnik, we find jokes funny, not just because they have surprising or incongruous endings, but rather because jokes make us aware of the flexible nature of our internal classification schemes in which such unexpected or incongruous fits are possible.

What generalizations, then, are emerging thus far about the relative efficacy of photos, line-drawings, cartoons and caricatures? Actually, the results are mixed. The work of Kennedy and Dwyer, support the view that outline pictures encode information about objects equally as well as, and in some cases better than, realistic photographs. On the other hand, the views of Perkins (1975), Arnheim and the Ryan and Schwartz experiment, while not supporting realism as the most informative mode, suggest an "ideal" or "schema" theory of recognition in which cartoons are more than just simplifications but mirror the organizing tendencies of the human mind, and, therefore, can be more quickly apprehended than realistic drawings or pictures. Hochberg has presented arguments which support the view that both cartoons and line-drawings will be more informative than highly realistic photographs.

It would be nice to stop here and say that, while it is not clear whether cartoons are more informative than line-drawings, at least we can rule out realism as being the best way to present pictorial information. However, a recent expriment by Perkins and Hagen (1980) forces us to hedge on making such a blanket statement. In their study, photographs were taken of 54 faces in both full and three-quarter profile views. A professional artist then made caricatures of each of the 54 faces in both views. (See Fig. 2.7). A subset of 15 faces were then chosen to form a training set. Some subjects saw caricatures of the 15 faces while others saw photos. The main question of the expriment was: Would seeing a caricature during training better enable subjects to sort the 15 faces from among the original set of 54 when presented together at a later time? The test faces could either be photos or caricatures in either full or profile views. It turned out that seeing caricatures does indeed permit better than chance recognition -- even when the test set is photos. However, subjects trained on photos did even better than the caricature group in identifying the training set from among the 54 test photos. These results then complicate the issue: they warn against making any pat generalizations about the relative informativeness of photographs versus caricatures and also call into question the "schema" theory of how our concepts (at least about faces) are stored. But the results of this experiment deserve to be followed up by further investigations. There may be alternative explanations of the poorer

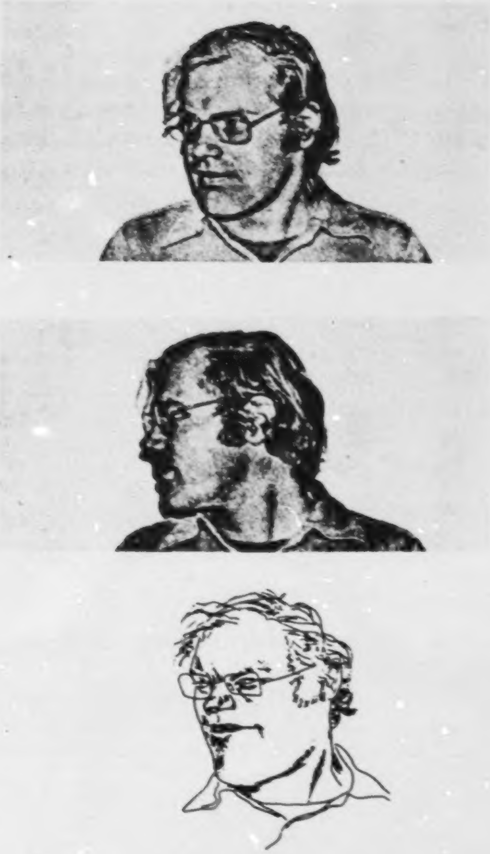


Figure 2.7. Sample photographs and caricatures used in the Perkins and Hagen (1980) experiment.

performance of the caricature group in this study. For example, although the caricatures were made by a professional artist, and do seem to capture the "key" features of the faces, notice that the style of these drawings are actually quite different than the style of the cartoon-hand in the Ryan and Schwartz study (Figure 2.4): the cartoon-hand uses bolder, simpler contours than do the caricatured faces in Figure 2.7 which use a finer, more detailed style of sketching, more reminiscent of the "shaded drawing" condition in the Ryan and Schwartz study. Thus, the question remains, at least for me: Would a bolder style of caricature, perhaps more akin to Perkins's drawings of Nixon in Figure 2.6, using simple contours and more exaggerated features, have facilatated later recognition in this experiment? Until attempts are made to systematically investigate such "style" issues, a definitive answer about the relative informativeness of photos and caricatures must be postponed.

2.4. Summary and Conclusion

The most common use of visual imagery on Telidon, at least initially, is likely to be what Arnheim refers to as their "picturing" function (see Chapter I) -- i.e., the use of an image to provide people information about the physical appearances of objects and events-- as opposed to, say, the "symbolic" function of imagery whose purpose is to give specific shape to abstract ideas (see Chapter III). In this chapter we examined this picturing function of visual images and explored some of its implications for Telidon. We can summarize the main themes of this chapter as follows:

(1) Learning to read pictures is not like learning to read words. Pictures are not arbitrary signs: pictures package information about the physical appearances of objects and events roughly the way real scenes do. There are, however, some interesting differences between pictures and real scenes which argue that some role must be granted to culture and convention in explaining how people adapt to the medium of pictures. In general, though, there should be no problem of pictorial "literacy" on Telidon: anyone who can see should be able to see a picture on Telidon.[1] But, being able to read a picture does not guarantee that we will read a picture in the exact way intended by its maker or even that two people will read a picture in the same way. Beyond a basic level of identification (e.g., it's a picture of snail), an image can give rise to many different propositions (fragile, slimy, slow, has antananae, etc.) depending on context.

(2) The ease and rapidity of reading a picture is not necessarily directly related to its degree of photographic realism. There are both theoretical and empirical grounds for believing that images which depart from mechanical likeness -- outline drawings, cartoons and caricatures -- may be more easily grasped by a

---

[1] A more exact statement should perhaps be that anyone who can "sense" can "sense" a picture. This is in light of Kennedy's (1974; 1978) innovative work on "picture books for the blind" where he shows that the congenitally blind are able to "see" raised line-drawings of familiar objects and spatial relations. This work also gives additional support to the view that the invariants of optical structure are amodal or abstract with respect to sensory modality and that these invariants are capturable by line-drawings.

human mind which builds its interpretations of visual inputs by fitting them to mental "prototypes" or "schemas," than are photographic images. In this chapter we explored the notion that cartoons and caricatures work so well because, in a sense, they mirror the "canonical" form of our mental schemas. A "schema" view of visual interpretation enabled us to formulate several hypotheses concerning how one could exploit Telidon's strength as a pictorial medium: i.e. its ability to generate clear, simple, line-drawings and graphics. Specifically:

(1) Can it be shown that outline-drawings of familiar objects are "equivalent" to high-fidelity photographs in so far as their ability to represent, on Telidon, most of the important variables of optical structure (e.g., shape, surfaces, edges, texture, cracks, etc.)?

(2) Despite a possible equivalence in terms of their ability to represent optical structure, are photographic quality images still preferred over realistic line-drawings by users (e.g. people may believe that photos don't lie)?

(3) How does the variable of transmission delay in constructing a pictorial image on Telidon affect user acceptance of the system? Namely, is there a "tradeoff" between the variables of degree of photographic realism, transmission delay and the "perceived objectiveness" of the image such that users are willing to sacrifice photographic quality images for outline pictures because of greatly reduced transmission delays in

picture construction? How is this tradeoff affected by characteristics of the retrieval task: e.g., "breadth first" versus "specialized" search?

These questions concerning the effects of delay, realism and type of informational retrieval task, appear amenable to empirical study and are a good example of how fundamental research into cognitive and perceptual processes could go hand-in-hand with the planning and evaluation of new visual telecommunications technologies.

CHAPTER III

IMAGES FOR THINKING: GRAPHIC SCHEMES FOR COMPLEX IDEAS

Simple graphic images and line-drawings can be used, not only as "pictures" in Arnheim's sense (see Chapter I) -- i.e., outlines which provide information about physical appearance -- but also as "symbols:" using a graphic image to represent concepts or ideas at a higher level of abstraction than the image itself. Simple graphics (e.g., a Venn diagram) can help us to understand complex ideas, especially in educational and scientific domains. It is here that an appropriate graphic schema, particularly when it exploits the image-transforming capabilities of computers, may outstrip verbal language alone as an aid to comprehension. (Note that it is not being claimed that a graphic schema by itself can supplant words. While language alone may not be the best way of getting across a scientific or technical concept, words will always be necessary to direct the user's processing of an abstract graphic schema). In this final part of the report I briefly consider a few examples of how simple graphic elements such as lines and geometric shapes can be effectively used in such activities as learning and thinking. What should be kept in mind in this section is that, while the graphic elements themselves may be quite simple, the underlying ideas they embody often require a great deal of specialized knowledge on the part of the user in order to be understood. This means that unlike outline-drawings or cartoons of objects and people, which should be easily identified by the average Telidon user, some graphic images will only be comprehended by a small subset of users with specialized needs in science, engineering, education, architecture, and so on.

3.0. Learning from Maps

Let us begin by looking at one of the most often cited applications of graphic imagery on Telidon: the use of maps. Maps are interesting because they seem to exist halfway between the pictorial and the symbolic. A map bears a certain degree of isomorphism to the territory it represents -- as in a typical street map. Such a map conserves information about some actual piece of territory: its shape, the geographic location of towns, cities, landmarks, and so on. Yet a map also contains certain conventional elements -- a black line can represent a highway, a dot an entire city, a cross a church, and so forth (see Bertin, 1973, for a detailed discussion of the "languages" of different kinds of maps).

Maps are useful precisely because they do not attempt to represent, in a one-to one correspondence, everything in the piece of territory, but are selective, choosing to represent only certain features of the territory. This selectivity makes it easier for user to perform certain calculations using the map; for example, to plot the quickest route between two cities. Interestingly, one can make an argument concerning the efficacy of maps which mirrors the arguments concerning the efficacy of pictures developed in the previous chapter. That is, the most effective maps may be those which distort objective realism in order to facilitate the calculation process. A good example suggested by Arnheim (1964) is that of a pocket subway map (see Figure 3.0) which eliminates geographic detail in order to accentuate the relevant topolo-

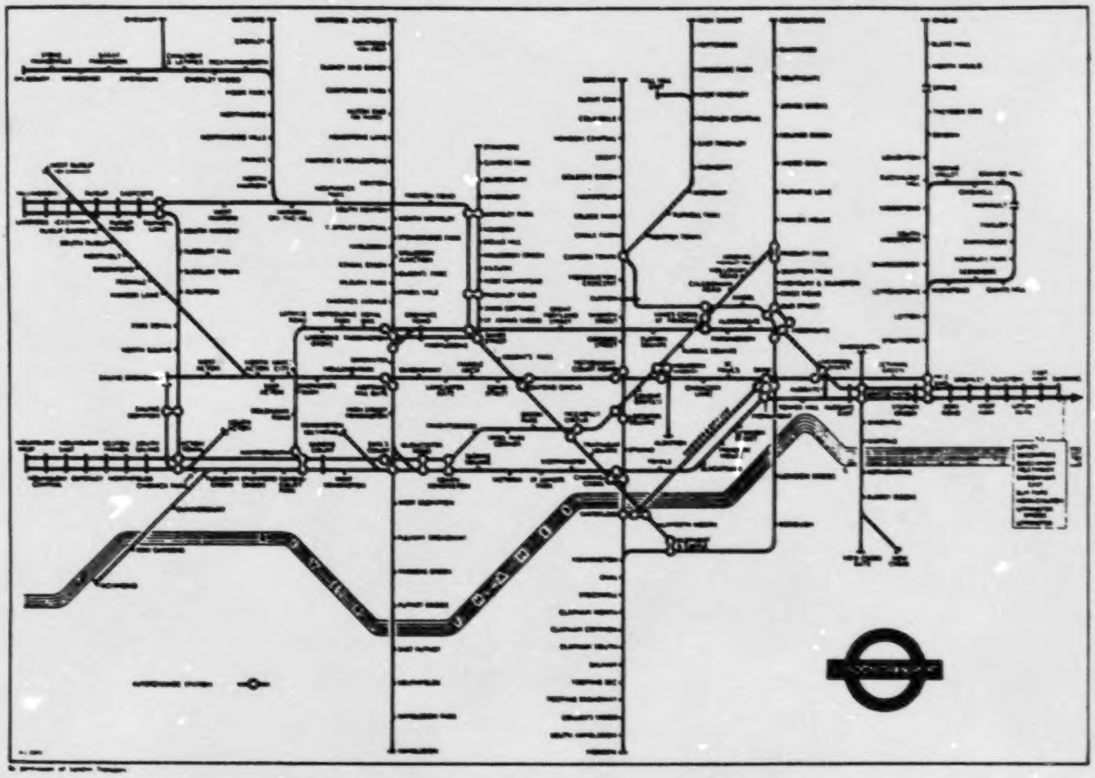


Figure 3.0. A subway map.

Source: Arnheim (1964)

gical relations between points: the subway map replaces routes with straight lines which intersect at either 90 or 45 degree angles: This makes it easier for the eye to grasp the sequence of stops and their interconnections. Schematic subway maps such as Figure 3.0 aid the human mind in its quest to organize incoming visual data.

The subway map designer's intuitions about how to achieve visual clarity by distorting reality have recently found support in psychological experiments where people were asked to reproduce a drawing or map of some familiar spatial layout such as the floor plan of an apartment (C.f. Norman and Rumelhart, 1975; Arnheim, 1975). These studies show that human memory is not geared to reproduce accurately spatial layouts, even of places with which one may be very familiar. Instead, people's maps drawn from memory often distort the shapes and interconnections between spaces, making them more straight and symmetrical than they really are, thereby serving to highlight functional, not physical reality. The map in Figure 3.1 (b), for example, shows a young girl's attempt to reproduce from memory the physical layout of her parent's apartment, the actual layout of which is given in Figure 3.1 (a). Her reproduction, like the subway map, reveals the tendency toward visual clarity by highlighting topological and functional connections between the spaces. The girl's map seems to be a better solution of how to represent the temporal, linear experience of moving through an environment -- movements where straight paths will be followed when possible -- than the actual floor plan which shows, simultaneously, the multiple connections.

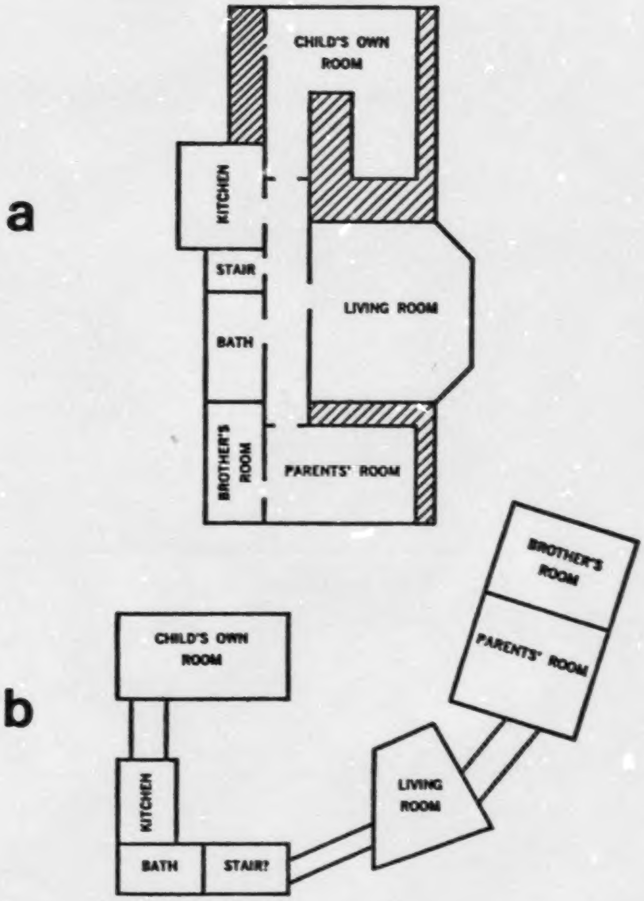

a
CHILD'S OWN ROOM
KITCHEN
STAIR
LIVING ROOM
BATH
BROTHER'S ROOM
PARENTS' ROOM
b
BROTHER'S ROOM
CHILD'S OWN ROOM
PARENTS' ROOM
LIVING ROOM
KITCHEN
BATH
STAIR?

Figure 3.1. The actual layout of an apartment (a) and a 12 year-old girl's reproduction from memory (b).

Source: Arnheim (1975)

3.0.1. Individual Differences in Learning Maps

There is little question that maps of many kinds, especially the simple topographic ones, are well within Telidon's graphics capabilities. But what is involved in acquiring knowledge from a map? Is reading a map as easy as reading a picture? Will most users be able to understand the information in maps on Telidon? Or are there individual differences in people's ability to read, remember, and use the information in maps? If there are differences, are they due to experience and therefore can be eliminated through training, or are the differences due to innate cognitive capacities? While not concerned with maps on Telidon, a recent study by Thorndyke and Stasz (1980) attempted to answer such questions.

In the first part of their experiment, the authors asked "experts" and "laymen" to attempt to learn the information in a ficticious town map well enough so that they could reproduce the map from memory and be able to solve navigational problems such as planning the best route between two points. The participants spoke aloud during the learning trials. A scoring system based on the accuracy of the reproduced maps permitted the authors to distinguish "good" learners from "poor" learners. They found that the good learners had strategies or "effective procedures" for learning spatial relations, for evaluating their own progress, and for focussing attention on unlearned information. The unexpected result of this experiment was that there was no correlation between being a "good learner" and prior expertise in using maps: What mattered was the use of "effective procedures."

In a follow-up experiment the authors wanted to know if teaching "effective procedures" to laymen of equal map learning ability would boost their performance compared with a control group instructed in "neutral" procedures. They found that, indeed, that one can teach people to be better map learners by teaching them effective procedures. However, the study also revealed another interesting finding: After the experiment, subjects were independently tested for "visual memory" ability -- the ability to accurately recall spatial information. An interaction was then found between training and visual memory. That is, people with "high visual memory" more easily learned the "effective procedures" than people with "low visual memory." In fact, subjects low in visual memory ability improved no more than the control subjects who were not instructed in "effective procedures."

This experiment raises some practical questions for the use of maps on Telidon. First, it implies that providing a map -- even a typical street map -- does not guarantee that it will be easily understandable by everyone (especially if one must remember the information in the map for later use). Maps, while built from simple graphic elements, can require complex information processing skills which may not be possessed by everyone in equal amounts. There appear to be clear individual differences in map learning. More seriously, however, is the question of the source of the individual differences: Are they due to lack of experience or innate cognitive capacities? If high visual memory is a prerequisite for learning to be a good map user, the teaching of effective procedures for helping people to learn from maps may be a waste of time -- unless one finds a teaching strategy for boosting visual memory (a seemingly innate skill). The wide-spread usefulness of maps, then, may be more limited than one might have originally suspected. In part, this may be because a map, like an architect's plan or blueprint, shows objects and layouts from top-views or other vantage points rarely encountered in normal everyday perceptual experience. But also, and equally as important, the goals of reading a map or blueprint, unlike that of reading a picture, is not merely one of identification, but usually involves using spatial information to perform some kind of problem-solving -- e.g., navigating between points, estimating distances, identifying distinctive land features, etc. And the ability to solve such problems may depend, not only of the visual clarity of the map, but also on the possession of specific information processing strategies seemingly rooted in fixed cognitive capacities. Clearly, more needs to be known about the cognitive building blocks involved in learning to use maps. Such research should provide more insight into the practical use of maps on information retrieval systems such as Telidon.

3.0.2. Movie Maps

The Thorndyke and Stasz experiment indicates that there may be individual differences, rooted in fixed cognitive capacities, in people's ability to acquire and use knowledge from maps. But is a "topographic" or "top-view" map the only means by which a user could acquire knowledge about a spatial environment? Perhaps by harnessing today's computer graphics technologies one could devise other kinds of maps which exploit people's everyday knowledge of the perceptual world. An innovative alternative to the ordinary street map, which brings together the technology of computer graphics with the optical videodisc, is currently being explored by the Architecture Machine Group at MIT. They call their project the "movie map" (Lippman, undated). The basic idea of the movie map is described by Lippman:

> ...At its simplest level, it (the movie map) may be regarded as a dynamic replacement for a topographic paper map: it can familiarize a user, or map reader, with a spatial environment. This familiarization, however, is accomplished by quite different means: the "map reader" explores space by participating in a simulated drive through it, seeing filmed sequences that replicate the actual views he would have were he in the space, driving. The experience of driving is made more intensive and involving through interaction: the user determines routes, turns, speeds, and points of view. He may also select the season, via a "season knob," and the visual mode of the tour: a photo, sketch, or animation. Thus the system does not simply repeat a guided tour, but allows a person to freely explore, at his own rate, and via his own path. (Lippman, page 4).

In addition to letting a user take a simulated drive through a spatial environment, the movie map system allows the user to retrieve ancillary data associated with buildings, streets, and so on that he may encounter during the drive. For example, a driver could stop off at the facade of a chinese restaurant and then request to see the menu. Thus a movie map can also be conceived of as a kind of dynamic iconic index.

The movie map works by storing previously filmed images of, say, a drive through a town (filmed from several angles simultaneously) on a pair of videodiscs -- each videodisc being able to store and provide random access to 54,000 television frames. Because the address of each frame can be coded into a form readable by a computer, programs can be developed which allow a user great flexibility in "editing" the coded frames: building sequences of images, determining travelling speeds, etc. The use of two videodiscs allows the user to execute a number of choices in travelling down a street; for example, deciding to make a left turn at an oncoming intersection. This is possible because while one videodisc is displaying its images simulating travel down one street, the second "look ahead" videodisc is automatically positioned to begin playing its own sequence of images of the street beginning at the intersection. When a turn is signaled by the user, the second disc begins playing its frames. Once the turn is completed, the first videodisc, which in the meantime has repositioned its own frames, begins to play so that the journey can continue uninterrupted at the next intersection.

The movie map contains many other innovative features (e.g., touch-sensitive signals for controlling the sequences, synthesized sound) but we do not have space in this report to cover them all. Before leaving the movie map however, we would like to question at least one basic assumption underlying it. Earlier we saw that the idea of the movie map was to be able to replace a topographic paper map of a region with a dynamic simulated drive through the region. This alternative was interesting in light of the potential problem of individual differences in map reading skills as suggested by the Thorndyke and Stasz study. That is, while some people may have trouble with a topographic map, anyone should be able to use a movie map since it simulates to a certain extent, what one would see actually by being there. But does a movie map really replace an ordinary street map? Are they really designed to fulfill the same function?

A pocket street map of a city, for example, while not showing the appearance of individual houses, buildings and landmarks, provides the user a global view of the territory: the spatial relations among the streets and landmarks. With such a map, it is possible to see where one is, where one wants to go, and to find the best way to get there. It displays the network of connections on a single spatial image. The street map does not tell you what the museum on Main and Park will look like once you arrive, however. How could you find out how to get to the museum on Park and Main, if you've never been there before, using the movie map as so far described? Only by actually experiencing the trip. And if the trip is complicated with many twists and turns, a single trip may not suffice because it would require holding in memory a great deal of sequential, temporal experience. A movie map then is linear and sequential in terms of experience of spatial relations among items as opposed to a topographic map where the multiple spatial connections are displayed simultaneously to the eye. For this reason a movie map, as currently described, would not be a good replacement for a topographic map. It would, however, be an excellent complement. In fact, the designers of the movie map have already anticipated this possibility by including computer animation effects which "generate an interactive overview of the town, and to remove excess detail that might confuse the user." The ability to alternate between map-like overviews and a realistic "drive" would seem to provide the best of both worlds. And, while such a system will probably not be available on a wide scale in the near future, the nature of the movie map is an exciting demonstration of why pictorial communications will certainly be an important part of computer mediated interactions in the future:

perhaps on some future version of Telidon.

3.1. Graphic Schemes for Thinking

A common experience for those who work in university or scientific settings is to enter a lecture hall or conference room and find the blackboard or flip chart filled with the previous occupants' graphic scribbles: boxes, arrows, circles, etc. Sometimes labels accompany these diagrams, sometimes they don't. An interesting game is to try to figure out what topic was being discussed based on the graphic images alone. Most of the time the graphic schemes will give some clues about the basic properties of the topic being discussed: although the exact topic may be difficult to identify. We usually recognize the scribbles as specifying "pure" relations without knowing exactly the nature of items being related; e.g., "overlap" (A Venn diagram); "flow" or "transmission" (a flow chart); "hierarchy" (a tree diagram), and so on. Such graphic schemas convey structural characteristics which can fit many different domains: a "flow" might occur in a computer program, a circuit diagram or a business organization. Why do teachers and scientists feel the need to make these graphic scribbles? If thinking, as some believe is "internal speech," why not simply say it in words alone? Do these simple boxes and arrows really aid thinking or are they merely attention grabbers or conventions, having no real connection with the abstract thoughts they are supposed to illustrate?

Psychologists have traditionally believed that thinking takes place either as "internal speech" or as "mental images" -- the latter being faint picture-like reproductions of perceptual experience (c.f. Paivio, 1971). While a treatise on the "language of thought" is beyond the mandate of this report (but see Fodor, 1980 for a recent discussion), we should point out that a few scholars have settled on a third hypothesis which may help us understand why "abstract graphic schema" aid thinking. Rudolph Arnheim, for example, has been a strong advocate of the notion that all thinking is "perceptual" as distinguished from "pictorial." According to Arnheim (1969), the medium of thought is neither words nor picture-like mental images, but some kind of abstract imagery which captures the relevant "generic" properties of objects, situations and events in a way which aids thinking about them. For Arnheim, the role of language in thinking is clearly subsidiary to the role of perceptual knowledge. In Arnheim's view, one simply cannot think in words; words are merely the pathways into the real medium in which thinking occurs -- the abstract imagery of generic shapes, relations and forces. The role of words is mainly to call forth the abstract perceptual knowledge underlying language and to activate the mental operations on this knowledge.

Obviously, it is impossible to directly observe the abstract images of thought, but we can get an idea of their attributes by looking at an experiment reported in Arnheim's book, Visual Thinking. In this experiment people were asked to invent graphic metaphors for "abstract" concepts such as "past, present and future," "democracy," and "good marriage/bad marriage." Although different subjects seldom invented the same graphic symbols for these concepts, they had no difficulty in finding some schematic image which captured, for them, certain aspects of these concepts. Figure 3.2, for example, shows one person's drawing for "good

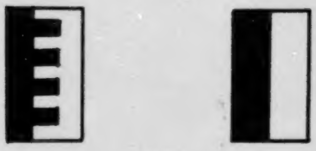


Figure 3.2. A visual metaphor for "good marriage/ bad marriage."

Source: Arnheim (1969)

marriage/bad marriage." What is interesting about this drawing is not that it "stands for" the meaning conveyed by the words "good marriage/ bad marriage" in any exact way. What is interesting is that despite the abstract and schematic nature of the images (there are no realistic pictures of wives, husbands, wedding cakes or flying rolling pins), we can recognize how the structural features of the shapes convey, on an abstract level, some important differences between what it means to have a good as opposed to a bad marriage. These differences might be paraphrased as follows: "In a good marriage there are two independent entities, yet they are interlocked to form a strong global unit; in a bad marriage, on the other hand, the two separate units do not mesh, but are joined, weakly, on a superficial level to form a weaker whole."[1] Again, let us emphasize that what is important here is not that a graphic schema can replace the meaning of complex concepts such as "marriage," "democracy" and so on, for which we have perfectly good words in english, but that abstract schematic forms can give visible shape to constellations of forces and relations (e.g., "interdependence") which can aid in thinking about these relations. I believe this is part of the reason that engineers, educators and scientists find graphic schemas so useful. In the same way that Figure 3.2 serves as a visual metaphor which helps structure our thinking about a concept such as "marriage," so a block diagram or schematic circuit can help structure our thinking about scientific or technical concepts by eliminating extraneous detail, and by highlighting important structural features and relations (containment, overlap, flows, etc.) not directly observable in the original phenomena.

---

[1] Our ability to comprehend such visual metaphors, despite the fact that the pictures and words share no common surface features, argues that the meeting ground between linguistic and perceptual knowledge is more abstract than either words or concrete realistic images. Similar views about the role of abstract schematic imagery in language and thought have recently been expressed by Verbrugge (1974), Lakoff (1980) and Mills (1980).

One common example of how a graphic schema gives visible shape to complex ideas is the "tree diagram." A tree diagram presents logical hierarchical connections among items to the eye in a way which outstrips the ability of language alone to transmit the same information. Imagine trying to piece together the links in a family tree or organizational chart from a verbal description alone (Johnson reports to Adams who reports to Clark who reports to Jenkins, who reports to... etc.). Another common example is the usefulness of computer flow charts whose visual structure help us "see" the interrelations amongst the parts of a complex program in a way that reading the instructions in the programming language would fail to do.

3.1.2. Dynamic Graphics

In the previous section we suggested that the best way to communicate a complex idea is not necessarily by using a complex graphic image. On the contrary, we suggested that the reason schematic graphics are so popular with scientists, educators and engineers, is not because they are complex, but because of their simplicity: they give shape to abstract properties and relations which are not directly observable in the original phenomena. Schematic diagrams of all sorts will be useful to people with specialized needs and are well within Telidon's current graphics capabilities. So far our examples of graphic schema have been restricted to the kind one might find on a blackboard or in a textbook: that is, static, frozen images. In this final section I will briefly consider the question of whether a still picture or graphic exploits the full potential of the computer (and therefore Telidon) as a tool for generating graphics which can help us think: What gains could be made in the comprehension of ideas by seeing images which move and change?

Problem-solving and learning always require going beyond the information given; that is, productive thinking usually involves a "restructuring" of one's initial perception of a problem, event or situation in order grasp the appropriate paths leading to a solution (Newell and Simon, 1972). Such "restructuring" operations are inherently dynamic; they involve changing or modifying one's initial internal model of a situation from one state of knowledge to a different state. And while a still image or graphic schema may facilitate restructuring operations, images which undergo transformations may do so even better. Arnheim's book on Visual Thinking provides an example. Consider the geometric figure in Figure 3.3. The problem is to find the area of the figure when lines A and B are given. Finding the solution to this problem may be difficult at first because of the perceptual system's tendency to see the geometric configuration as the simplest possible structure: a square overlaid with an oblique parallelogram. However, the solution is more easily perceived if the lines are seen as representing another kind of

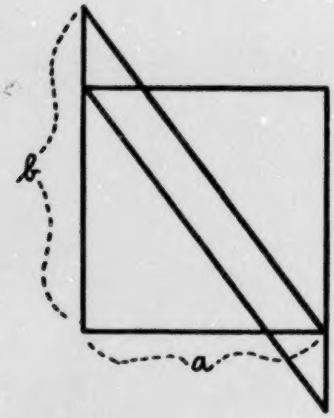


Figure 3.3. Geometric problem-solving.

Source: Arnheim (1969)

figure: namely, two overlapping right-triangles each with the area of $\frac{1}{2}(A \cdot B)$. Although the required "restructuring" can be activated by the verbal suggestion alone, it would be much more compelling if the mode of presentation were an animated film: one observes the figure slowly pulling apart revealing the two separate triangles which move about as separate entities against the background. Then, the two triangles move together once again, and overlap to form the still-image shown in Figure 3.3. The creation of such dynamic sequences to help teach mathematical or scientific concepts is possible, not only using traditional techniques of film animation, but is also well within the graphics capabilities of modern computers. Moreover, the computer potentially allows, not only dynamic imagery, but also can invite the user to interact with it; to make choices and to perceive the consequences of those choices (e.g., Lutz and Rigney, 1976, showed the superiority of interactive computer graphics over purely verbal modes of presentation in teaching elementary concepts in electrochemistry).

One of the most ambitious attempts to evaluate the communicative potential of dynamic graphics is the book by William Huggins and Doris Entwisle called Iconic Communications (1974). These authors suggest that, with the advent of the computer as a tool for storing and manipulating images, the time is ripe for a science of "iconics" which would bring together knowledge from education, linguistics, communications, psychology and computer science in order to harness the production of dynamic imagery for educational purposes. While, in my opinion, the authors tend to fall into the trap of wanting to substitute images for words (see Chapter I of this report), they do provide useful examples and report some experiments which shed light on how dynamic graphics can help understanding in education. Their book also contains a large annotated bibliography.

One particularly relevant example is the attempt by the authors to develop a notation of "dynamic symbols" -- arbitrary iconic signs which change state -- in order to teach some basic notions of flow in an electric circuit: How some elements are active (batteries) while other elements are passive (resistors); how the quantity of flow "divides" when more than one path is available: how the topological arrangement of elements influences flow (parallel versus sequential arrangements of elements). The challenge here, as the authors point out, is that one cannot simply film the appearance of a battery with some resistors; While a battery and a resistor may look different, the physical differences do not help one understand their electrical properties. Moreover, the authors point out that the traditional ways of portraying such differences, by using mechanical pointers or digital voltmeters, for example, do not exploit the potential of iconics for showing structural and causal relations.

The solution to these problems was to produce a film in which the components of an electrical circuit and their functional properties were represented by "dynamic symbols." Figure 3.4 shows several frames from this movie. Note that the authors have abandoned the standard way of representing circuits in textbooks for a simpler notation which they hoped would more naturally convey the essential ideas. Instead of the

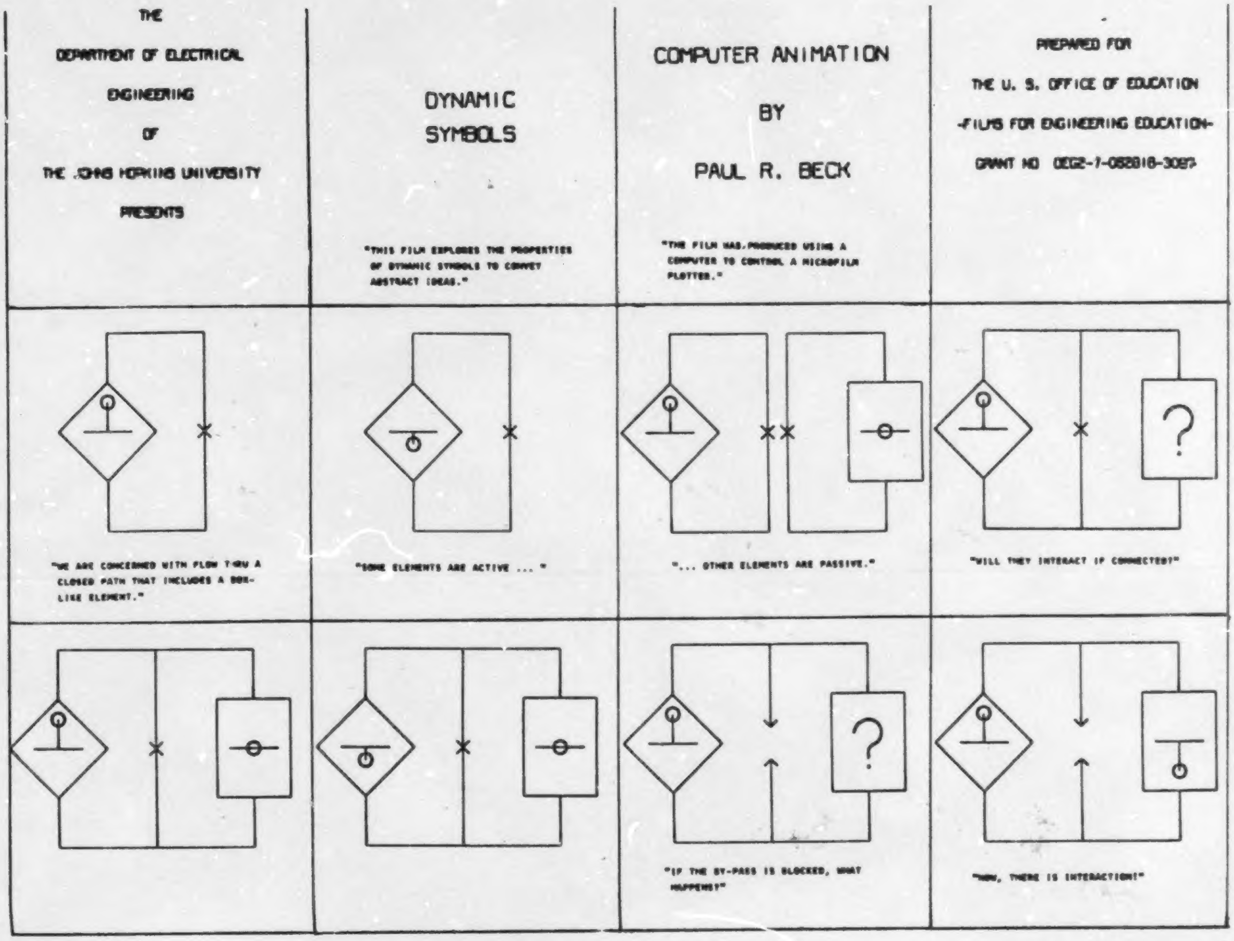

THE
DEPARTMENT OF ELECTRICAL
ENGINEERING
OF
THE JOHNS HOPKINS UNIVERSITY
PRESENTS
DYNAMIC
SYMBOLS
"THIS FILM EXPLORES THE PROPERTIES OF DYNAMIC SYMBOLS TO CONVEY ABSTRACT IDEAS."
COMPUTER ANIMATION
BY
PAUL R. BECK
"THE FILM WAS PRODUCED USING A COMPUTER TO CONTROL A MICROFILM PLOTTER."
PREPARED FOR
THE U. S. OFFICE OF EDUCATION
-FILMS FOR ENGINEERING EDUCATION-
GRANT NO OEG2-7-062818-3097
"WE ARE CONCERNED WITH FLOW THRU A CLOSED PATH THAT INCLUDES A BOX-LIKE ELEMENT."
"SOME ELEMENTS ARE ACTIVE ... "
"... OTHER ELEMENTS ARE PASSIVE."
"WILL THEY INTERACT IF CONNECTED?"
?
?
"IF THE BY-PASS IS BLOCKED, WHAT HAPPENS?"
"NOW, THERE IS INTERACTION!"

Figure 3.4. Still-frames from the "dynamic symbols" film.

Source: Huggins and Entwistle (1974)

usual symbol for a voltage meter, for example, they have included a "lollipop" symbol -- a small circle whose length visibly increases or shrinks from a reference line depending on changes in the amount of current flowing through the circuit. By visibly showing the changes to such simple graphic elements, the film tries to communicate basic electrical concepts of flow, resistance, additivity and so on. In a follow-up test of the effectiveness of the "dynamic symbols" film, the authors showed that both engineering and non-science students, after seeing the film, performed better than a control group in grasping basic electrical concepts. There was one interesting difference between the groups, however. The engineering students interpreted the film in the language of circuit theory while the non-science students "anthropomorphized" the ideas (in somewhat sexist fashion): i.e., describing a battery component as a "boy" (active) trying to communicate with his "girl" (the passive resistor). Seeing the dynamic symbols as visual metaphors for social action attests, once again, to the power of abstract graphics to express general, structural characteristics which can be instantiated in many different domains: i.e. the cartoon enables one to see common patterns across scientific and social phenomena. Finally, it is worth noting with the authors that, although the simple animated graphics appeared to be a success, such films, while not difficult to produce on a technical level, are difficult conceptually. It takes quite a bit of time and talent to translate conceptual ideas, especially scientific ones, into the language of dynamic graphics. Obviously, these constraints of visual imagination would apply to future uses of abstract graphics for educational purposes using Telidon.

3.1.3. Motion versus Sequence

One final note before closing the discussion of dynamic graphics. There is little question that seeing an image or graphic undergoing change can help the restructuring operations necessary for productive thought. And communicating changes and transformations can be accomplished through animation sequences in which still images succeed each other at a rate which gives the viewer the illusion of smooth "apparent" motion. (It is currently possible to create continuous motion sequences on Telidon, although the process of creating apparent movement with the current page creation system would appear to be somewhat laborious). In film and television the illusion of continuous movement is accomplished by transmitting 24 or 30 still-picture frames every second. But is "apparent" motion always necessary to understand dynamic events? Not necessarily. It is useful, in my opinion, to distinguish between the perception of "motion" and the comprehension of "sequence." To understand the meaning of a dynamic sequence requires recognizing the state changes or transformations which specify particular kinds of events (walking, melting, colliding, breaking, etc.). The specification of such transformations may not require the illusion of continuous motion, but only the ability to integrate information from several successive samples of still pictures, each of which shows a discrete state of the event. An ordinary comic-strip does exactly that. For example, it is not necessary to perceive continuous motion in order to understand that the cartoon-strip in Figure 1.1 uses four discrete frames to specify the notion of a dynamic event; i.e., an object "looming" toward an observer. What is required is knowledge of the "syntax" of cartoon-strips. Similarly, the "dynamic symbols" film described above could easily be understood as a cartoon-strip or even a "slide-show" where successive still images replace each other on the user's screen.

Making the distinction between the mind's ability to understand "sequence" and the perceptual system's ability to perceive "apparent motion" suggests that one could explore methods of presenting users information about dynamic visual events based on "sampled" still-frames. Such techniques could greatly conserve transmission bandwidth as well as save valuable memory space in the case of digital encoding of dynamic visual sequences (see Mills, 1972 for an early discussion of this question).

3.2. Summary and Conclusion

One of the main themes emerging throughout this paper is that, to be useful to the human mind, an image does not have to be a faithful replica of reality. In this chapter we examined some ways in which graphic schemas -- highly abstract images using simple graphic elements -- could be used to help us understand complex concepts, including the layout of spatial environments and the teaching of mathematical and scientific ideas. We can summarize the main points of this chapter as follows:

(1) One of the most often cited applications of graphic imagery is the use of maps. We found that the most effective maps may not be the most realistic, but are those which actually "distort" reality by eliminating information and by visually clarifying the topological and functional connections among geographical entities. (e.g., a pocket subway map).

(2) Despite their obvious usefulness, there may be individual differences in people's ability to use and learn from maps. Good map learners use "effective procedures" (information-processing strategies) in learning maps and in problem-solving with maps, that poor learners do not use. Moreover, the ability to learn "effective procedures" may be linked to fixed cognitive capacities such as visual memory ability. The important point for Telidon is that, unlike cartoons, caricatures or outline pictures of objects, which should be easily understood by everyone, certain kinds of pictorial representations, which may have previously been thought easily understood by everyone, may actually be harder to use, for some people, than previously suspected (e.g. maps, blueprints, plans, etc.). These kinds of graphics may require complex information-processing skills which some people may not possess and, worse, may find extremely difficult to acquire.

(3) If some people have trouble with a topographic map, everyone should be able to understand a "movie-map" -- a simulated, user-controlled "drive" through a town via videodiscs and computer graphics technology. While an impressive demonstration of a possible application of pictorial communications for information-retrieval (great for shut-ins, for example), we questionned whether a movie map could replace the good old street map. The street map and movie-map seem to call upon different modes of information-processing and are intended to fulfill different

needs. Combining a movie-map with a topographic map would provide the best of both worlds: but one wonders if such systems will be available, for all but very specialized user populations, in the near term.

(4) Simple graphic schemas such as boxes, arrows, circles, can help thinking about complex concepts in scientific, educational and technical areas. It was argued that the usefulness of graphic schemas was not arbitrary, not due to convention or habit -- but exploits the human mind's need to seek out underlying simplicities and to focus on generic forces and relations: to grasp the structural characteristics of phenomena not usually visible, in any simple way, in the phenomena themselves.

(5) While graphic schema help thinking, dynamic graphic images -- images which undergo transformations -- may be even more help in learning and problem-solving. This is because a dynamic graphic can help one perform the mental "restructuring" operations (the transforming of one's initial internal model of a problem) so that the appropriate solution will be more readily grasped. We reviewed some examples of how computer generated dynamic symbols were being used to teach concepts in electrical theory.

(6) Finally, we suggested that while dynamic graphics can help thinking, it is useful to distinguish between "motion" and "sequence." The mind can grasp "sequence" by integrating successive discrete samples of still images as specifying the transformations characteristic of dynamic events, without having to per-

ceive the illusion of continuous motion in these transformations as in movies and television. The distinction between "motion" and "sequence" suggests that techniques should be investigated for providing users information about dynamic visual sequences as successive still images. Such "still frame" techniques would afford savings of transmission bandwidth and computer memory.

SUMMARY AND CONCLUSIONS

This report began by asking two broad questions: (1) What can theories of perception and cognition tell us about how people understand and use pictures? and (2) What implications does this research have for the use of pictures on Telidon? In trying to answer these questions we have covered a lot of ground. Yet, the territory of pictorial communications is so vast that many relevant theories and experiments could not be included in a single report. However, a supplementary bibliography is provided in Appendix B for the interested reader.

Let us summarize, then, the main insights into the role of pictures on Telidon which have emerged from our analysis and review of the a literature:

(1) Visual images of various kinds can potentially play an important role on future videotex systems, including Telidon. But the use of a picture or graphic will not automatically boost the usefulness and acceptability of Telidon. The use of a picture brings with it certain costs to information-providers, system operators and to users. In particular, page-designers should beware of the ad-hoc use of visual imagery on Telidon which makes processing demands on the user, but contributes nothing, or worse, which distracts from the overall comprehension of the ideas.

(2) Information-providers and page-designers should not try to stage a contest between pictures and words, i.e., put great effort in thinking of ways to replace clear verbal descriptions

with visual images. While pictures of various sorts will certainly be useful on Telidon, it is unlikely that a purely iconic language can be invented to supplant the descriptive functions of language. The real challenge for IP's is to understand how pictures can serve as a "conceptual base" which facilitates verbal comprehension, and how language guides the interpretation of pictures in various contexts.

(3) A picture can give rise to different propositions by being placed in different contexts. The problem of the multiple meanings of pictures is a direct result of the flexibility and context-sensitivity of the human conceptual system. This is why pictures present difficult challenges to applied areas of research relevant to Telidon: e.g., the automatic classification and retrieval of images and the development of "iconic indexes." Research into the nature of human categorization, especially the prototypicality of concepts, is pertinent to the problem of developing verbal and pictorial indexes on Telidon. It was argued that the best pictorial indexes would be built out of simple, abstract, graphic elements yet be recognizable as pictorial representations of particular objects and events. Good examples are the iconic symbols used during the Montreal olympics.

(4) For uses of imagery as "pictures" -- images which provide information about the physical appearance of objects and events -- there should be no problem of "pictorial literacy" on Telidon. Anyone who can see the visual world, should be able to see the visual world through a picture. However, pictures do not provide

information to the eyes exactly as do real-scenes, but demand certain kinds of adaptation. Being able to "read" a picture, though, does not mean we will necessarily read off of a picture the exact proposition intended by its maker or that everyone will read the same meaning off of a picture. The effect of context on meaning still holds even for realistic pictures.

(5) The pictures which are the easiest and quickest to read on Telidon may not be those with the highest degree of photographic realism. There is some theoretical and empirical research suggesting that outline-drawings are able to conserve most of the important variables of optical structure. Moreover, cartoons and caricatures may be more easily recognized than photographs and realistic line-drawings. This may be because, in making a cartoon or caricature, an artist has managed to give external form to the schematic or "canonical" nature of our internal codings of perceptual objects and events. These hypotheses have practical implications for Telidon which excels as a medium for generating high-quality, schematic line-graphics. But more research needs to be carried out on the need for pictorial realism on Telidon which takes into account possible trade-offs between the degree of photographic realism, the nature of the retrieval task, transmission delays and people's beliefs in the "veridicality" of different kinds of visual representation.

(6) Simple schematic graphics can play an important role on Telidon, especially in scientific and educational domains, where a graphic

can help the user understand abstract concepts and principles, not directly visible in the surface features of phenomena. While graphic images can aid thinking, graphic images which show transformations, may be even more effective because they help one perform the "restructuring" operations necessary to productive thought. A problem for future research would be to devise efficient techniques on Telidon for generating dynamic graphics.[1] It was pointed out that people may be able to comprehend successive still-picture sequences of visual events without needing to perceive the illusion of continuous motion. This distinction between "motion" and "sequence" suggests that techniques could be developed to generate dynamic visual sequences on Telidon in a way that conserves bandwidth and computer memory.

(7) Page designers should be aware that, unlike the "picturing" function of imagery, the "symbolic" use of schematic imagery to represent abstract ideas -- while easy to generate through the PDI's -- will require specialized knowledge to be interpreted and therefore not necessarily be easily understood by the average Teldon user.

(8) Graphic imagery such as maps, blueprints and plans are fruitful candidates for research on Telidon. Such representations exist halfway between the symbolic and pictorial functions of imagery and may be more difficult to understand and use than previously thought. In fact, there is some evidence that there are significant individual differences in people's ability to use and

---

[1] The Telidon research group has already created a graphics communication protocol called Picture Manipulation Instructions (PMI's) which among other things allows for dynamic graphics.

learn from maps, and that these differences are rooted in innate cognitive capacities.

In closing, we would like to point out what we believe is one of the important side-benefits, from the viewpoint of the behavioural scientist, of doing applied research on new telecommunications systems such as Telidon; namely, theories of cognition, perception and communication can not only teach us something about Telidon, but Telidon can help one identify new theoretical and empirical issues. We saw, for example, that in order to evaluate the need for pictorial realism on Telidon, more basic research should be carried out on the "informational equivalence" of different kinds of pictorial representations which takes into account such variables as type of retrieval task, beliefs in pictorial veridicality and transmission time. Studying the links among such variables would not only help one make practical decisions about the use of pictures on Telidon, but would tell us more about the principles underlying the human visual system. The important point is to acknowledge that the question of "What are the implications?" can mean, not only what are the implications of theories for Telidon, but also what are the implications of Telidon for the development of theories of human information-processing?[1]

---

[1] Telidon raises basic questions not only about pictures, but about other fundamental aspects of the human information processing system. To take just one example, the Department of Communications' Behavioural Research Group, in studying the optimal methods of indexing and data-base organization, are also coming to grips with basic questions of "cognitive mapping" -- i.e. finding out how our internal maps are structured can help one derive better indexing schemes. See Telidon Behavioural Research 1, Ottawa, February, 1980, for a review of the projects and problems being investigated by the Behavioural Research Group.

REFERENCES

Arnheim, R. The Dynamics of Architectural Form. Berkeley: University of California Press, 1975.

_________. Art and Visual Perception. Berkeley and Los Angeles: University of California Press, 1964.

_________. Visual Thinking. Berkeley: University of California Press, 1969.

Attneave, F. "Representation of Physical Space." In A. Melton and E. Martins (Eds.), Coding Processes in Human Memory. Washington. D.C.: V.W. Winstone & Sons, 1972.

Bartlett, F.C. Remembering. Cambridge: Cambridge University Press, 1932.

Bellugi, U. and Klima, E. The First Signs of Language. Cambridge: Harvard University Press, in press.

Bertin, J. Sémologie graphique: les diagrammes, les réseaux, les cartes. Paris: Mouton, 1973.

Bliss, C.K. Semantography: Bliss Symbolics. Semantography Publications, 1965.

Bower, T.G. R. "Discrimination of Depth in Premotor Infants," Psychonomic Science, 1964, 1, 368.

_________. "Slant Perception and Shape Constancy in Infants," Science, 1966, 151, 832-834.

Bower, G.H.; Karlin, M.B., & Dueck, A. "Comprehension and Memory for Pictures." Memory & Cognition, 1975, Vol. 3(2), 216-220.

Bown, H.G.; O'Brien, C.D.; Sawchuck, W., and Storey, J.R. Picture Description Instructions (PDI) for the Telidon Videotex System. Canadian Department of Communications Technical Note. No. 699-E, November, 1979.

_________. A General Description of Telidon: A Canadian Proposal for Videotex Systems. Department of Communications Technical Note. No. 697-E, December, 1978.

Bransford, J.D., & Johnson, M.R. "Considerations of some Problems of Comprehension." In W.G. Chase (Ed.), Visual Information Processing. New York: Academic Press, 1973.

Bregman, A.S. "Perception and Behavior as Compositions of Ideals." Cognitive Psychology, 1977, 9, 250-292.

________. "Asking the 'What for' Question in Auditory Perception." In Kubovy, M. and Pomerantz, J.R. (eds.) Perceptual Organizations. Hillsdale, New Jersey: Lawrence Erlbaum, (in press).

Carmichael, L.; Hogan, H.P. & Walter, A.A. "An Experimental Study of the Effect of Language on the Reproduction of Visual Perceived Form." Journal of Experimental Psychology, 1932, 15, 73-86.

Chomsky, N. Aspects of the Theory of Syntax. Cambridge: MIT Press, 1965.

Clowes, M.B. "Transformational Grammars and the Organization of Pictures. In Grasselli (Ed.) Automatic Interpretation and Classification of Images. New York: Academic Press, 1969.

Duncan, H.F.; Gourlay, N. & Hudson, W. A Study of Pictorial Perception among Bantu and White Primary School Children in South Africa. Johannesberg: Witwatersrand University Press, 1973.

Dwyer, F.M. A Guide for Improving Visualized Instruction. State College, Pa.: Learning Services, 1972.

________. Strategies for Improving Visual Learning. State College. Pa.: Learning Services, 1978.

Fischler, M.A. "Machine Perception and Description of Pictorial Data." Proceedings of the International Joint Conference on Artificial Intelligence. Washington, D.C., 1969.

Firschein, O. and Fischler, M.A. "A Study in Descriptive Representation of Pictorial Data." Pattern Recognition. Vol. 4, 361-377.

Fodor, J. The Language of Thought. Cambridge: Harvard University Press, 1975.

Gibson, J.J. "A Theory of Pictorial Perception." Audio-Visual Communications Review, 1954, 1, 3-23.

________. The Senses Considered as Perceptual Systems. Boston: Houton-Mifflin, 1966.

________. "The Information Available in Pictures." Leonardo, 1971, 4, 27-35.

Gombrich, E.H. The Visual Image. Scientific American. September, 1972.

________. Art and Illusion. New York: Pantheon, 1960.

Goodman, N. Languages of Art: An Approach to a Theory of Symbols. Indianapolis: Bobbs-Merrill, 1968.

Goodman, N. "On J.J. Gibson's New Perspective." Leonardo, 1971, 4, 359-360.

Gopnik, A. "Did you Hear the One about the...: A Meta-Cognitive Theory of Humour." Unpublished paper. McGill University, 1979.

Hagen, M. & Jones, R. "Cultural Effects on Pictorial Perception: How many Words is one Picture Really Worth?" In R. Walk and H. Pick (Eds.), Perception and Experience. New York: Plenum, 1978.

Hebb, D.O. The Organization of Behavior. New York: Wiley, 1949.

Hochberg, J. & Brooks, V. "Pictorial Recognition as an Unlearned Ability," American Journal of Psychology, 1964, 75, 624-628.

Hochberg, J. The Representation of Things and People. In Gombrich, E.H., Hochberg, J. and Black, M. Art, Perception and Reality. Baltimore: Johns Hopkins University Press, 1972.

__________. In the Mind's Eye. In Haber, R.H. (Ed.). Contemporary Theory and Research in Visual Perception. New York: Holt, Rinehart and Winston, 1968.

Huggins, W.H. and Entwisle, D.R. Iconic Communication: An Annotated Bibliography. Baltimore: Johns Hopkins Press, 1974.

__________. Iconic Communications. IEEE Transactions on Education, 14, 1969, 158-163.

Jahoda, G. & McGurk, H. "Pictorial Depth Perception in Scottish and Ghanaian Children: A Critique of Some Findings with the Hudson Test." International Journal of Psychology, 1974, 9(4), 225-267.

Johnson, A.R. Dialogue and the Exploration of Context: Properties of an Adequate Interface. Presented at the Fourth Annual International Symposium of the American Society of Cybernetics. Washington, D.C. 1970.

Kennedy, J.M. A Psychology of Picture Perception. San Francisco: Jossey-Bass, 1974.

Kennedy, J.M. & Fox, N. "Pictures to See and Pictures to Touch." In D. Perkins and B. Leandar (Eds.). The Arts and Cognition. Baltimore: Johns-Hopkins Press, 1977.

Kolers, P.A. Some Formal Characteristics of Pictograms. American Scientist, 57, 1969, 348-363.

Lakoff, G. Images Schemas. Paper presented at the Second Annual Meeting of the Cognitive Science Society, Yale University, New Haven, 1980.

Lippman, A. Movie-Maps: An Application of the Optical Videodisc to Computer Graphics. Paper published by the Architecture Machine Group, Massachusetts Institute of Technology, undated.

Loftus, E.F. & Palmer, J.C. Reconstruction of Automobile Destruction: An Example of the Interaction Between Language and Memory. J.V.L.V.B., 1974, 13, 585-589.

Mead, M. and Modley, R. Communication among all People, Everywhere. Natural History, 77, 1968, 56-63.

Mills, M.I. The Effectiveness of Real-Time versus Still-Frame Television Channels for Interactive Problem-Solving. Systems Engineering Report 145. Bell Northern Research, Ottawa, Ontario, 1972.

__________. On Fitting Cartoon-Strips to Descriptions: A Study of some Relations Between Perception and Language. Unpublished Doctoral Dissertation, McGill University, 1980a.

__________. Cognition and the Comprehension of Cartoon-Strip Analogies. Paper presented at the National Symposium for Research in Art: Representation and Metaphor. University of Illinois, 1980b.

Mills, M.I. and Bregman, A.S. On Fitting Cartoons to Descriptions: Viewpoint Transformations in Event Comprehension. Unpublished, 1980.

Minsky, M. A Framework for Representing Knowledge. In P.H. Winston (Ed.) The Psychology of Computer Vision. McGraw-Hill, 1975.

Modley, R. Pictograms and Graphs. Harper and Brothers, 1947.

Muter, P. The Role Graphics in Learning on Videotex: Evidence from other Media. Technical Memorandum, Department of Communications, Ottawa, Ontario, 1980.

Narasimhan, R. On the Description, Generation, and Recognition of Classes of Pictures. In Graselli, A. (Ed.) Automatic Interpretation and Classification of Images. New York: Academic Press, 1969.

Newell, A. & Simon, H.A. Human Problem Solving. Englewood-Cliffs, New Jersey: Prentice-Hall, 1972.

Newport, E.L., and Bellugi, U. Linguistic Expression of Category Levels in a Visual-Gestural Language: A Flower is a Flower is a Flower. In Rosch, E. and Lloyd, B. (Eds.) Cognition and Categorization, New York: Lawrence Erlbaum, 1978.

Norman, D.A. & Rumelhart, D.E. Explorations in Cognition. San Francisco: W.H. Freeman & Sons, 1975.

Olson, R.K.; Pearl, M.; Mayfield, N. & Millar, D. "Sensitivity to Pictorial Shape Perspective in 5-year Old Children and Adults." Perception and Psychophysics, 1976, 20(3), 173-178.

Paivio, A. Imagery and Verbal Process. New York: Holt, Rinehart and Winston, 1971.

Perkins, D.N. "Caricature and Recognition." Studies in the Anthropology of Visual Communication. Spring 1975, 2(1).

__________. Pictures and the Real-Thing. Paper presented at the Conference on Processing of Visible Language at Niagara-on-the Lake, September, 1979.

Perkins, D.N. & Hagen, M. "Convention, Context and Caricature." In M. Hagen (Ed.), The Perception of Pictures. Vol. 1: Alberti's Window: The Projective Model of Pictorial Information. New York: Academic Press, 1980.

Piaget, J. La psychologie de l'intelligence. (1947). Translated as the Psychology of Intelligence. London: Routledge and Kegan Paul, 1950.

Rigney, J.W. and Lutz, K.A. The Effects of Interactive Graphics Analogies on Recall of Concepts in Science. Technical Report, No. 79, Department of Psychology, University of Southern California, 1976.

Rock, I. An Introduction to Perception. New York: MacMillan, 1975.

Rosch, E. Cognitive Reference Points. Cognitive Psychology, 1975. 7, 532-547.

Rosch, E. & Lloyd, B. (Eds.) Cognition and Categorization. New York: Lawrence Erlbaum, 1978.

Ryan, T.A., & Schwartz, C.B. "Speed of Perception as a Function of Mode of Representation." American Journal of Psychology, 1956, 69, 60-69.

Shaw, R. and Bransford, J.D. Perceiving, Acting and Knowing: Toward an Ecological Psychology. N.J.: Lawrence Erlbaum, 1977.

Thorndyke, P.W., and Stasz, C. Individual Differences in Procedures for Knowledge Acquisition from Maps. Cognitive Psychology, 12, 137-175, 1980.

Winograd, T. Understanding Natural Language. N.Y.: Academic Press, 1972.

Wollen, K.A.; Weber, A. & Lowry, D.H. Bizareness versus Interaction of Mental Images as Determinants of Learning. Cognitive Psychology, 1972, 2, 518-523.

Woolfe, R. Videotex: The New Television/Telephone Information Services. London: Heyden & Son, 1980.

Verbrugge, R.R. The Comprehension of Analogy. Unpublished Doctoral Dissertation. University of Minnesota, 1974.

APPENDIX A

AN INTRODUCTION TO THE PSYCHOLOGY OF PICTURE PERCEPTION

by
Adam Gopnik

Rather than attempting a survey of unrelated titles, the following bibliographies are grouped by theme and subject, with relevant titles listed together, and a general description of the debate which centers on these works provided below. We believe that this approach provides a better picture (so to speak) of the kinds of recurring issues and positions that have characterized this fluid and always controversial discipline, than could be provided by a mere list.

A.- The Classical Psychological Theory of Picture Perception

1. Dergowski, J. "Illusion and Culture" in R. Gregory and E. Gombrich (Eds.) Illusion in art and nature. New York: Charles Scribner's Sons, 1973.
2. Gombrich, E.H. Art and Illusion. Princeton: Princeton University Press, 1960.
2. " " Meditations on a Hobby Horse. New York: Phaidon, 1963.
3. " " "The What and the How: Perspective Representation and the Phenomenal World" in Logic and Art, Essays in Honour of Nelson Goodman, ed. Richard Rudner and Israel Schleffer. Indianapolis: Bobbs-Merill, 1972.
4. " " "The Mask and the Face" in Art, Perception, and Reality, Maurice Mandelbaum (Ed.). Baltimore: Johns Hopkins Press, 1972.
5. Goodman, N. Languages of Art: An Approach to the Theory of Symbols. Indianapolis: Bobbs-Merrill, 1968.
6. Hochberg, J. "Nativism and Empiricism in Perception." In Psychology in the Making. L. Postman (ed.) New York: Holt, Rinehart, and Winston, 1962.
7. Silvers, Anita. "Show and Tell" in The Arts, Cognition and Basic Skills. St-Louis: CEMREL, 1978.
8. Wilson, B. and Wilson M. "Recycling Symbols" in The Arts, Cognition and Basic Skills. St-Louis: CEMREL, 1978.

Consider a perspective drawing of a cube. Why do we see it as a cube, rather than as a peculiarly shaped two dimensional array? If, as in fact is true, an infinite number of different arrangments will produce the same picture (our "cube" might be a figure with the "front" and "side" both on the same plane, and an awkwardly shaped triangle pasted in above, for instance), if we could logically interpret this figure as any one of number of infinite figures, why in fact does our eye see one and only one figure -- a cube?

Classical psychological theory, at least from the time of Bishop Berkley, has asserted that it is experience that makes us see the drawing as a cube. We see what we expect to see. We see a cube because in our experience of the world most angles are right angles and parallel lines converge. By perceptual habit, we see the object that in normal experience would most likely be producing the array of lines and shapes that reaches our eye.

In 1960, E.H. Gombrich took this simple piece of classical psychology and gave it a radical re-interpretation that, it may be said, engendered a new discipline. Gombrich proposed that all representation was shaped by convention, by experience codified in the form of artistic tradition. We produce pictures, Gombrich suggested, not by copying what we see, but by manipulating conventional representational stereotypes, what Gombrich called "schemata." This did not mean that all representation involved the simple recycling of conventional symbols -- on the contrary, the whole history of post-renaissance art was the story of artists' attempt to alter or "correct" these conventional schemata in the pursuit of more and more accurate representations. But even the most "photographic" depiction would always be marked by the employment of convention, of those formulas for depiction that sum up our knowledge of the perceptual world.

While Gombrich always emphasized the pull of the conventional stereotype, the tendency to represent the unfamiliar in terms of the familiar (so that a drawing of an unfamiliar thing -- a new animal or an unusual landscape has in the history of art always been rendered in terms of some familiar schema), he also emphasized the way in which the artist could correct or change these schemata in the direction of greater and greater representational accuracy. Gombrich believes in the possibility of an "objective" naturalism -- although he also believed that verisimilitude would only be achieved by the slow evolution of convention and modification, of schema and correction. But a significant school of theorists -- chiefly philosophers -- whose work followed Gombrich's, believed that naturalism was itself a figment, and that Gombrich's "schemata" were simply arbitrary symbols for things in the world, and that the question of their resemblance to the things they claim to represent was irrelevant. Most prominent among these theorists is Nelson Goodman, who in his famous The Languages of Art argued that the conventional schemata of Western Art are like words in a spoken language, with a purely arbitrary and assigned relationship to the things in the world they stand for. To Goodman, pictorial representation "depends not upon imitation or illusion or information, but upon inculcation. Almost any picture may represent almost anything; that is, given picture and object there is usually a system of representation... under which the picture represents the object and there are usually many such systems" (Goodman, 1968 above, p. 38). Goodman believes that almost anything can stand for almost anything else. Goodman believes in the hold of convention and experience, but unlike Gombrich he divorces our experience of the visual world from our experience of the pictorial world -- they are different worlds, Goodman believes, which provide us with essentially different kinds of experience. We understand pictures not in terms of our experience of the world, but in terms of our experience of other pictures. We see a cube, to return to our example, not because a cube is the most familiar form of the infinite possible forms that this drawing could represent, but because we know that in our culture this is what a representation of a cube looks like - in our system of conventions, this is the symbol that stands for a cube.

In recent years, variations of Goodman's hypothesis have achieved a certain vogue in writings on the subject (surprisingly, Gombrich's position is often conflated with Goodman's, although they are in fact in complete opposition). Philosophers are especially fond of this position, because it centers debate on the question of how signs and symbols work, a traditional philosophical question, and avoids the question of whether a drawing looks like what it is supposed to represent, a question which bores philosophers. Anita Silvers has recently further pursued the notion that pictures are symbols. Silvers believes that even if a picture can be said in some sense to "resemble" something in the world, the way we use pictures is not to represent things, but to denote, to exemplify, to express. Brent and Marjorie Wilson have pursued a similar line of argument; in their article "Recycling symbols" they claim that the pictorial language consists of a vocabulary of conventional symbols, and that when we want to draw an unfamiliar thing we "map" the new thing onto a pre-existent symbol -- we render the unfamiliar in terms of the familiar. This is like Gombrich's position, with the exception that Gombrich believes that conventional symbols can be "corrected" in the direction of greater naturalism, while the Wilson's, following Goodman, believe that conventional symbols can only be "recycled," placed in some new context or combined with some other symbol that allows them to stand for some new thing. The Wilson's point out, for instance, that children learn to tell stories with pictures not by learning to represent their experience but rather by learning a certain set of conventional narrative devices (scene changes, birds-eye views, transitions from "long shots" to "close-ups," etc.). The analogy, again, is always with spoken language -- Goodman et al. believe that just because we can encode information about the world in pictures does not mean pictures resemble the world, anymore than the fact that we can encode information about the world in English means that the English language "resembles" the world.

Among the empirical support offered in support of this position, emphasis is often placed on the work of the anthropologist Dergowski, who argues that what is perceived as a good representation differs from culture to culture (and from age to age, so that children and adults have different ideas about what constitutes a good picture of an elephant say) -- which suggests that to some degree the interpretation of pictures is a learned skill. Kennedy however, (see below) has disputed this interpretation of Dergowski's research.

Julian Hochberg, whose original research we will discuss in greater detail below, has outlined the intellectual background of this school of thought very well in his "Nativism and Empiricism in Perception."

What can we learn from these studies? Gombrich's theories have already begun to have an impact in education, where a reaction has set in against the notion that copying "corrupts" the eye of the child, and the idea has begun to be accepted that learning conventional schemata through copying is an important part of attaining visual literacy. And the question of whether the language of pictures is socially tempered or universal deserves more investigation, and may have important ramifications as any graphics system seeks a world-wide audience.

B.- The Gestalt Theory

1. Arnheim, Rudolph. Art and Visual Perception: A Psychology of the Creative Eye. Berkley: University of California Press, 1958.
2. Arnheim, Rudolph. Towards a Psychology of Art. Berkley: University of California Press, 1967.
3. " " "Introduction: Psychology and the Arts" in Perception and Pictorial Representation, Calvin F. Nodine (ed.) St-Louis: CEMREL, 1979.
4. Bernheimer, R. The Nature of Representation. New York: N.Y.U. Press, 1961.

Consider our drawing of a cube once again. We know that there are similar figures, called reversing or ambiguous figures, in which the faces of a drawing of a cube can reverse -- what at first appears to be the front plane becomes the back plane. What is important is that we can only see it one way at a time -- we cannot keep both figures in view at the same time. This suggests to "gestalt" psychologists that there are some underlying fundamental rules that govern our perception of pictures. They propose that we see the drawing of a cube as a cube not because that is the most familiar way to see it, but rather because that is the simplest way to see it. Familiarity with the world, they argue, does not always help us in reading pictures: it is perfectly possible to conceal very familiar shapes, like letters, by embedding them in lines and flourishes that have never been seen before. We interpret pictures not through the probabilities of experience, but through underlying organizational rules in our mind.

Rudolf Arnheim has extended this approach, and through it developed an important and complex psychology of art. To a gestalt psychologist like Arnheim, what is crucial is that we do not perceive pictures by passively recording their information and subsequently "interpreting" it. Instead we impose structure on the world in order to make it meaningful. Pictures, in Arnheim's theory, are most successful when their intrinsic structure corresponds to the structure of relationships in the thing they represent. To take a simple example: When we look at a landscape, we know that an artist can never actually "match" or "imitate" the actual brightness of reflected light in the natural world. What he can match is the relationships among lights -- he can get the relative brightness right and so create a convincing likeness. It is not the matching of individual components that matters in creating a good likeness, but the depiction of the similarity in their pattern of relationships. Or to take another example, why do some still pictures of galloping horses look inert, while others convey the "feeling" of movement quite successfully? Arnheim suggests that it is because the successful pictures capture in the structural relationships of their shapes and lines the inherent structural qualities of dynamism and movement. In response to the classical school, Arnheim asks how conventional representational sterotypes ever were invented in the first place? He suggests that it is because these conventional "symbols" successfully capture the crucial structural qualities of the things they depict -- they depict what is central to our mind, the structural relationship of the whole, and ignore insignificant and transient detail. A stick-figure of a man, for instance, is a good representation because it captures the crucial structural qualities of the thing it intends to depict -- we accept it as a representation not through social tradition but because its form satisfies our innate psychological need for parsimony and simplicity in descriptions.

Arnheim's approach has been called "non-cognitive," because it puts the emphasis on the structure of the picture itself, rather than on subsequent inferences we might make about the picture in our mind. But Arnheim's most profound belief is in the potential of pictures to represent any kind of knowledge or feeling. He believes that there are kinds of knowledge that can *only* be represented visually. If we want to understand a chemical synthesis, for instance, while we can get some grasp of it through words and numbers, we can really only effectively get a sense of it through a picture that can represent the impact and outflow of dynamic forces. Arnheim believes that a study of the underlying structuring principles of the mind can enable graphic artists to make more effective and visually intelligible displays (for an example of this, see his introduction to *Perception and Pictorial Representation*). It remains to be seen whether Arnheim's insights can be codified into a set of procedures, however. Even at this preliminary stage, though, we can learn from Arnheim that we need not be unadventurous in the kinds of things we attempt to use pictures for. It should be noted, however, that Arnheim's theory lacks a mechanism, an hypothesis to explain why some visual patterns have the expressive structure he claims for them. He can say, for instance, that a circle intrinsically embodies the qualities of "feed-back" or "homeostasis," but in doing so he is perhaps really only engaging in the kind of impressionistic analysis that is as old as art criticism, albeit this time with an impressive sounding vocabulary borrowed from psychology. One should be warned that the gap between what Arnheim promises and what he delivers can sometimes be rather great -- he is not above the occasional platitude (The Napoleonic stance of the actor suggests pride ("embodies the structural qualities of pride") because the head is raised above the level of normal human interaction, and proud people always want to lord it over the rest of us).

The non-cognitive aspect of Arnheim's work, its emphasis on an analysis of the structure of the picture, had a profound effect on the dominant school of the moment in the study of picture perception; that school which takes its inspiration from the work of J.J. Gibson and which centers on the concept of "information."

C.- Information Theory and Picture Perception

1. Gibson, J.J. "A Theory of Pictorial Perception" in Audio-Visual Communication Review, 1954, Vol. k, 3-23.
2. " " "The Information available in Pictures" in Leonardo, 1971, no. 4, 27-35.
3. Hagen, M.A. "Picture Perception: Towards a Theoretical Model." In Psychological Bulletin, 1974, 81, 471-479.
4. Kennedy, J.M. A Theory of Picture Perception. San Francisco: Josey Bass Inc. 1974.
5. " " "Icons and Information" in D. Olson (Ed.) National Society for the Study of Education Yearbook, 1976.
6. Pirenne, M.H. Optics, Painting and Photography. Cambridge, Cambridge University Press, 1970.
7. Ward, J. "A Piece of the Action" in Perception and Pictorial Representation. Calvin F. Nodine (ed.). St-Louis, CEMREL, 1979.

Where previous psychologists had been impressed by the seeming ambiguity of our visual world, and had therefore concluded that our perception of that world must be constructed by mental inferences guided by experience, J.J. Gibson, with a certain common sense clarity, emphasized how much information there is in the visual world. Gibson emphasized what most psychologists had overlooked, that as we move through the visual world most of the ambiguities are resolved -- we discover what is invariant in the visual array, and we discover the sources of the apparent ambiguity which appear in a fixed and static array. Gibson pointed out that most psychologists had missed the boat by studying visual perception only in fixed and static array, when in fact it is characterized by constant feedback through a continuous, seamless flow of information over time.

Gibson encountered a predictable difficulty in dealing with pictures. For a picture, by its very nature, is static and fixed -- we can't walk around a Van Gogh hospital room. How then, do we understand pictures? Gibson first proposed (1954) that a picture successfully depicted things by reproducing the array of light we would observe if we were looking at the real thing. But a little reflection shows that very, very few pictures are like this -- we understand outline drawings and cartoons that don't bear the least resemblance to the point-by-point array of light in the thing they represent. So Gibson amended his theory of picture perception (1971) and suggested that what pictures reproduce is the information in the original array - information about boundaries, contours, texture gradients, etc.

J.M. Kennedy developed this idea and has tried to make it as specific and precise as possible. Kennedy has attempted to show that because one can represent the information in the visual world with drawn outlines, the language of outlines is therefore a kind of lingua franca for representation. According to Kennedy everyone, no matter what their cultural background, can understand the language of outline, and you can depict just about anything in the world with it. Kennedy shows, following Gibson, that this is not because we learn a set of conventions but because the language of outline reproduces the crucial information about the visual world. What really matters in vision are discontinuities, boundaries, contours, places where one texture replaces another, or where one shape occludes another. Since lines can depict the basic elements, the significant information, in the visual environment, lines therefore have the capacity to depict anything that is visible.

We can see that while the Gibson-Kennedy position is much like Arnheim's in its emphasis on the actual structure of the picture, apart from the learning we use to understand it: but it is essentially unlike Arnheim's in its treatment of the problem of abstraction. For Arnheim, an abstract representation is often a better depiction than a realistic one because it reproduces the salient structural relationships that are stored in the mind. For Gibson and Kennedy, abstraction may be an aid to greater clarity in reproducing information, like dying a slide to make the significant parts more evident, but they are much more committed to the notion that a good picture should successfully reproduce not an abstract structure of relationships but the "given" information in the light.

Most battles of the moment in the realm of picture perception are fought in these lists. Pirenne has attempted to combine the notions of abstract structure with the notion of reproducing information by speaking of "canonical form," the conceptual form of things which exists in our mind and guides our perception of the world, in an ongoing cybernetic process. Other theorists have attempted to extend Kennedy's approach in the solution of other problems in pictorial perception. Ward, for instance, has looked at the problem of "apparent" movement in pictures. In opposition to Arnheim, who suggests that it is the picture's ability to reproduce the abstract structure of dynamism which makes it able to convey a sense of movement, Ward suggests that it is the picture's reproduction of significant informationin the world that makes it successful. Arnheim says that a good depiction of a galloping horse will depart considerably from natural appearances in order to emphasize the abstract structure of movement through the use of an intrinsically dynamic pattern of shapes and lines. Ward supposes that a good depiction of a galloping horse will be one in which the actual information in the world is reproduced with the greatest clarity. What is at stake here is the role of abstraction and distortion in picture-making: do extensive departures from natural appearances aid in reproducing essential structural relationships, or are the most successful pictures those which ingeniously reproduce the existing information in the visual world?

As we will see below, these issues are fundamental to our understanding of the nature and effectiveness of cartooning and caricature. But even if we question this aspect of Kennedy's research, there is still much that we can learn from him. Kennedy emphasizes that outline is the most efficacious graphic tool ever invented, and his research suggests that the most successful graphic technology will be that which allows us to produce the richest outline drawings. Kennedy's work would suggest that "sketch-pad" and related technologies which can produce smooth and continuous contours will be much more effective than technologies which produce drawings through filling in blocks of colour (and, indeed, we have noted informally that subjects almost always use the "block" technologies to attempt to create contours).

Kennedy's current work investigates the capacity of the blind to understand pictures.[1] He suggests that the blind can understand outline representations through a tactile mode. This suggests that all the information we receive from the sensory world may be encoded in such a way that it is accessible to the language of outline, that is, to representational systems that emphasize contours and discontinuities, even where this language is haptic rather than visual. Conceivably, television graphics might play some role in this kind of work with the handicapped.

D.- The Physiology of Picture Perception

1. Berlyne, D.E. "Ends and Means of Experimental Aesthetics" in *Canadian Journal of Psychology*. 1972, 26, 303-325.
2. " " *Studies in the New Experimental Aesthetics*. New York: Wiley, 1974.
3. Gopnik, A. and Mills, M.I., Interpreting Single Pictures as Dynamic Events. Unpublished, 1980.
4. Hochberg, Julian. *Art and Perception*. New York: Academic Press, 1979.
5. " " "Visual Art and the Structures of the Mind' in *The Arts, Cognition and Basic Skills*. Stanley Majeda, ed. St-Louis: CEMREL, 1979.
6. " " "The Representation of Things and People" in *Art, Perception and Reality*. Maurice Mandelbaum (ed.). Baltimore: Johns Hopkins Press, 1970.
7. " " "Some of the Things that Pictures are" in C.F. Nodine (ed.). *Perception and Pictorial Representation*. St-Louis: CEMREL, 1979.

Along with the growth of work which attempts to analyze the kind of information available in a picture, work has also been done recently which attempts to show how the physiology of vision of the eye and brain can affect our perception of pictures. While the work of Julian Hochberg transcends categorization, it is perhaps most interesting when Hochberg demonstrates how relatively simple constraints in the eye and brain can profoundly affect our reading of pictures. In "Some of the Things that Pictures Are," for instance, Hochberg tries to show how the physiological distinction between foveal and peripheal vision in the eye has been exploited by artists. The eye registers fine detail only within the very small foveal region at the center of the visual field, while the periphery registers only gross and generalized form. In an impressionist picture, Hochberg argues, the effect of an initial or "fresh" look is recreated by reproducing a scene as it is seen by the periphery of the eye, where our first impressions of a new scene are in fact registered. Hochberg explains the success of simplified outline drawings and caricatures along similar lines - since these forms are recognizable further out into the field of peripheral vision, the eye has to do less work than it does when examining a detailed reproduction, where each part of the picture has to come under the scrutiny of the fovea. We are "pre-wired," so to speak, to operate most efficiently with generalized, non-detailed forms. Hochberg's work on the pictorial consequences of the existence of these two modes of seeing could well have a profound impact on the design of pictures and texts for optimal efficiency of perception.

As we will see in our review of texts in applied psychology, the success of a representation cannot be separated from its aesthetic appeal. Berlyne has attempted to understand our aesthetic preference for certain patterns in terms of a model of "cortical arousal." This system, Berlyne supposes, increases its response to a visual pattern in accord with the complexity and novelty of the pattern. If the pattern is too novel or complex, however, the response drops off. The most aesthetically successful patterns are those which achieve a balance between novelty and familiarity. This is not exactly news to aesthetic philosophers - it is in fact the idea that underlies the most ancient "canons" of beauty - but it is original in its attempt to explain these preferences in terms of physiology, in terms of stimulus habituation and fatigue.

E.- *Using Pictures to Teach Things*.

1. Dwyer, F.M. *A Guide for Improving Visualized Instruction*. State College, Pennsylvania, Learning Services, 1972.
2. " " "Improving Visuals for Televised Instruction" in *Improving College and Universitu Teaching*, 18: 288-291
3. Klahr, D. (Ed.) *Cognition and Instruction*. Hillsdale: N.J.: Lawrence Erlbaum Associates, 1976.
4. Perkins, D. and Leondar, F. *The Arts and Cognition*. Baltimore: Johns Hopkins Press, 1977.
5. Perkins, D. "Talk about Art" in S.Majeda (ed.) *Art and Aesthetics: An Agenda for the Future*. St-Louis, Missouri. CEMREL INC, 1977.
6. " " " A Definition of Caricature and Caricature and Recognition" in *Studies in the Anthropology of Visual Communication*, 1975, II, 1.
7. Ryan, T.A. and Schwartz, C. "Speed of Perception as a Function of the Mode of Representation." *American Journal of Psychology*, 1956, 69, 60-69.

With the spread of interest in audio-visual aids in schools in the late nineteen fifties, researchers began to study whether or not pictures did in fact aid in learning - and to study how, if they did, they could be used most effectively.

The study of Ryan and Schwartz (1956) indicated that departures from photographic fidelity in drawings tended to aid their educational effectiveness. Ryan and Schwartz made various kinds of pictures of the same objects - black and white photographs, ink and wash "modelled" drawings, outline drawings, and cartoons. The difference in style between outline drawings and cartoons in this experiment is significant; the outline drawings, while they eliminated a great deal of detail and obviously did not attempt to reproduce the "array of light," did attempt to reproduce the actual contours of the depicted object. In the cartoons, on the other hand, "cartoon conventions" were employed - the Figures had the kind of smooth, overly-curved contours that we are familiar with from Disney and his followers, with the cartoon's usual distortion of relative proportion.

Ryan and Schwartz discovered that cartoons needed less exposure in order to be recognized subsequently by subjects than any other mode of drawing. The photographs and high-fidelity line drawings needed the longest times. Clearly, some distortions -- those used in caricatures and cartoons - can aid perception more than strict adherence to natural appearances.

Francis Dwyer pursued experiments in a similar vein for the next fifteen years. Dwyer found that realistic photographs of organs added nothing to students understanding of a medical lecture, while outline drawings of the same organs contributed significantly. In Dwyer's work, cartoons, in the sense that Disney or Ryan and Schwartz employed them, were not tested. It may well be that they would have been more effective still.

Dwyer also discovered that the use of colour, even where it seemed to add no particular information, significantly aided learning, suggesting that aesthetic considerations may be very important even in the most seemingly "prosaic" employment of pictures. An attractive picture is much more useful than an ugly one. (We have noted in Berlyne's work an attempted theoretical explanation of this educational fact). Dwyer also discovered that people expected highly realistic, detailed drawings to be the most effective learning aids - although in fact they are the least effective. This suggests a deep seated, initial resistance to the employment of "schematic" graphics in educational environments - a resistance which any new graphics system will have to overcome.

Dwyer looked at the particular problems that television presented in educational contexts. He discovered that line drawings were once again most effective (again, cartoons in the conventional sense were not tested) and that the amount of detail worked against the effectiveness of graphics on television. At the same time he discovered that increasing the size of the televised image decreased its effectiveness as an educational aid. Perhaps Hochberg's distinction between foveal and peripheal vision (above) can help explain this fact.

More recently, David Perkins has looked afresh at the problem of cartoons and caricature and their extraordinary effectiveness in aiding memory and recall tasks. Perkins had subjects trace photographs of well known people - politicians especially. These drawings were mostly unrecognizable. Then he asked his subjects to exaggerate features of the sitter. The second set of drawings were much more recognizable than the first - besides being much funnier.

As we can see, these experiments certainly falsify any theory of picture perception which emphasizes point-by-point fidelity to the information in the light as the source of a picture's success as a representation. Kennedy has suggested that cartoons and caricatures work simply by exaggerating the language of line-drawings. But a careful investigation may well reveal that there is something more involved than simple geometric enlargement of significant detail. Hochberg has suggested that caricatures work because they correspond more closely to the "canonic image" of things, the schematicized memory image we hold in our minds. Recently, Gopnik and Mills (undated) have suggested that the effectiveness of cartoon drawings often depends on their ability to summarize a visual sequence that in ordinary perception takes place over time (the expression of a face, the action or potential action of a figure). Are cartoons a simple extension of the language of outline drawing, or do they speak a different language? It is a question that demands further research.

APPENDIX B

SUPPLEMENTARY BIBLIOGRAPHY

Anderson, John R. and Paulson, Rebecca. "Interference in Memory for Pictorial Information." Cognitive Psychology, Vol. 10(a), April, 1978, 178-202.

Antes, James R. "Components of Pictorial Informativeness." Perceptual and Motor Skills. Vol. 47 (2), 459-464.

Bagget, Patricia. "Memory for Explicit and Implicit Information in Picture Stories." Journal of Verbal Learning and Verbal Behaviour. Vol. 14 (5), October, 1975, 538-548.

Bahrick, Harry P. and Gharrity, Kathleen. "Interaction amont Pictorial Components in the Recall of Picture Captions." Journal of Experimental Psychology: Human Learning and Memory. Vol. 2, March, 1976, 103-111.

Blake, Tyler. "Motion in Instructional Media: Some Subject-Display Mode Interactions." Perceptual and Motor Skills. Vol. 44 (3), June, 1977, 975-985.

Bruner, J.S. and Mackworth, N.H. "How Adults and Children Search and Recognize Pictures." Human Development, Vol. 13, 1970, 149-177.

Cornell, Samuel D. "Effects on Retention of Redundancy and Simultaneity of Aural and Visual Messages." Dissertations Abstracts International. Vol. 36 (6B), December, 1975, 2997-2998.

Clark, Robert C. "An Experimental Study of the Effects of Five Audio-Visual Presentations Upon Mental Imagery and Learning." D.A.I. Vol. 37, December, 1976, 3044-3045.

Dallenbach, Robert B. "A Study of Verbal and Imaginal Memory of Children and Adults Using Human Faces, Objects and Words in Paired Associate Recognition Tasks." D.A.I., 1974, 3985.

Ferguson, Eugene S. "The Mind's Eye: Nonverbal Thought in Technology." Science, Vol. 197, August, 1977, 827-836.

Fleming, Malcolm, C. "The Picture in your Mind." A.V. Communications Review, Vol. 25, 1977, 43-62.

Franken, R.E. and Davis, J. "Predicting Memory for Pictures from Rankings of Interestingness, Pleasingness, Complexity, Figure/Ground and Clarity." Perceptual and Motor Skills. Vol. 41, August, 1975,

Franken, R.E. and Rowland, G.L. "Organizational Factors in Picture Recognition Memory." Acta Psychologica. Vol. 42, November, 1978, 441-452.

Fredrico, E.A. and Montague, W. "Recognition Memory as a Function of Encoding Strategy and Stimulus Codability." Journal of Experimental Psychology. Vol. 104, November, 1975, 660-668.

Friedman, S.L. and Stevenson, M. "Perception of Movement in Pictures." In What Then, are Pictures?: The Psychology of Representational Art. M.A. Hagen (ed.) (in press).

Hagen, M.A. and Elliot, H.B. "An Investigation of the Relationship Between Viewing Condition and Preference for True and Modified Linear Perspective." Journal of Experimental Psychology: Human Perception and Performance. 1976, 479-490.

Honeck, R.P. "Five Questions on Semantics." Bulletin de Psychologie. 1976, 116-124.

Hunter, Patricia A. "Mental Imagery and Learning in a Community College: A Study to Determine if the Ability to Image Mentally Affects Learning from Pictorial or Verbal Presentations." D.A.I. Vol. 37, December, 1976. 3031-3032.

Just, Marcel A. and Carpenter, Patricia A. "The Semantics of Locative Information in Pictures and Mental Images." British Journal of Psychology. Vol. 66, November, 1975, 427-441.

Kosslyn, Stephen M. and Locklear, Eileen D. "Childrens Drawings as Data about Internal Representations." Journal of Experimental Child Psychology. Vol. 23, April, 1977, 191-211.

Kolers, P.A. "Some Psychological Aspects of Pattern Recognition" In Recognizing Patterns, P.A. Kolers and M. Eden (ed.). Cambridge: M.I.T. Press, 1968.

Kolers, P.A. "Reading Pictures and Readings Texts." In The Arts and Cognition. Baltimore: Johns Hopkins Press, 1977.

Kriska, Steven. "Representing Multivariate Data Points as Schematic Faces." D.A.I. Vol. 37, November, 1976, 2545-2546.

Long, Robert E. "The Effects of Imagery Stimulation in the Motor Performance of Blind and Sighted Subjects." Dissertations Abstracts International. Vol. 35, December, 1974, 3554-3555.

Marks, L.E. "On Coloured-Hearing Synesthesia: Cross-Modal Translations of Sensory Dimensions." Psychological Bulletin, Vol. 82, May, 1975, 303-331.

Mandler, Jean M. and Johnson, Nancy S. "Some of the Thousand Words a Picture is Worth." Journal of Experimental Psychology, September, 1976, 529-540.

Moles, Abraham A. "Image and Visual Communication - Towards a Methodology." Communications. Vol. 2, 1976, 330-348.

Nelson, Douglas R.; Reed, Valerie S and Walling, John R. "Pictorial Superiority Effect." Journal of Experimental Psychology: Human Learning and Memory. Vol. 2, September, 1976, 523-528.

Nelson, Thomas M. "Enjoyment of Perspective Transformations Induced by Movement." Canadian Journal of Behavioural Sciences. Vol. 8, July, 1976, 287-297.

Neisser, U. and Kerr, N. "Spatial and Mnemonic Properties of Visual Images." Cognitive Psychology, Vol. 5, 1973, 138-150.

Paivio, Allan. "Perceptual Comparisons Through the Mind's Eye." Memory and Cognition. Vol. 3, November, 1975, 635-647.

Paivio, Allan. Imagery and Verbal Processes. Toronto: Hold, Rinehart and Winston, 1971.

Pellegrino, J.W.; Siegel, A.W. and Dhawan, M. "Differential Distraction Effects Through Short Term and Long Term Retention of Pictures and Words." Journal of Experimental Psychology: Human Learning and Memory. Vol. 2 (5), September, 1976, 541-547.

Pickford, R.W. Psychology and Visual Aesthetics. London: Hutchison Educational Service. 1972.

Pittenger, John B. and Shaw, Robert E. "Aging Faces as Viscal Elastic Events: Implications for a Theory of Non-Rigid Shape Perception." Journal of Experimental Psychology: Human Perception and Performance. Vol. 104, November, 1975, 374-382.

Potter, Mary C. "Short Term Conceptual Memory for Pictures." Journal of Experimental Psychology: Human Learning and Memory. Vol. 2, September, 1976, 509-522.

Potter, Mary C. and Faulconer, B.A. "Time to Understand Pictures and Words." Nature, Vol. 253, February, 1975, 437-438.

Rigney, Joseph W. and Lutz, Kathy A. "Effect of Graphic Analogies of Concepts in Chemistry on Learning and Attitudes." Journal of Educational Psychology. Vol. 68, June, 1976, 305-311.

Scull, L. Robert. "Instructional Effectiveness of Two Types of Visual Illustration in Facilitating Specific Geographic Learning at the College Level." Dissertations Abstracts International. Vol. 35, May, 1975, 7038-7039.

Shaver, Phillip,; Pierson, Lee and Lang, Stephen. "Converging Evidence for the Functional Significance of Imagery in Problem Solving." Cognition: International Journal of Cognitive Psychology. Vol. 3, 1974-75, 359-375.

Shepherd, R.N. and Podgorny, P. "Cognitive Processes that Resemble Perceptual Processes." In Handbook of Learning and Cognitive Processes. W.K. Estes, (ed.) Hillsdale, N.J.: Erlbaum Associates, 1977.

Shepherd, R.N. and Chipman, S. "Second Order Isomorphism of Internal Representations: Shapes of States." Cognitive Psychology. 1970, 1-11.

Shumway, Helen. "An Experimental Study in the Teaching of Visual Structural Unity by the Use of Film." Dissertation Abstracts International. Vol. 35, March, 1975, 5791-5792.

Standing, L. and Smith, P. "Verbal-Pictorial Transformations in Recognition Memory." Canadian Journal of Psychology. Vol. 29, December, 1975, 316-326.

Snodgrass, Joan G. and Assiaghi, Anthony. "The Pictorial Superiority Effect in Recognition Memory." Bulletin of the Psychonomic Society. Vol. 100, July, 1977, 1-4.

Sullivan, A.M. "The Relative Effectiveness of Instructional Television." Interchange 1976-77. Vol. 7, 46-51.

Vitz, Paul C. "Visual Science and Modernist Art" In Perception and Pictorial Representation. Calvin F. Nodine (ed.) New York: Praeger, 1979.

Walk, Richard D. "Perception of Imagery: An Empirical Intergration." Bulletin de Psychologie. Vol. 30, May-June, 1977, 737-738.

Yunurmey, A. Daniel. "Hypermesia for Pictures but nor for Concrete or Abstract Words." Bulletin of Psychonomic Society. Vol. 8. August, 1976, 115-117.